Multivariate Approaches to School Climate Factors and School Outcomes

Diego Alonso Carrasco Ogaz

PhD in Psychology

University of Sussex

Thesis submitted by Diego Alonso Carrasco Ogaz to the University of Sussex for the qualification of Doctor of Philosophy in Psychology

April 2016

UNIVERSITY OF SUSSEX

DOCTOR OF PHILOSOPHY IN PSYCHOLOGY

MULTIVARIATE APPROACHES TO SCHOOL CLIMATE FACTORS AND SCHOOL OUTCOMES

Summary

School climate is a crucial concept used to explain school differences. Nevertheless, this concept is elusive in the literature, conveying different meanings. To address the relation between school climate and school outcomes, its historical roots are reviewed and a multivariate approach to it is proposed, in contrast to a unidimensional conception. In four papers, this strategy is used to study associations among various school climate factors (SCFs) and school outcomes, including teacher turnover, teacher job satisfaction, students' math achievement, and students' social attitudes.

In paper 1, schools serving more socioeconomically disadvantaged students are found to present higher rates of teacher turnover. A complementary study shows that SCFs (supportive school leadership, positive school relationships, and academic monitoring) present differing effects on teacher turnover.

In paper 2, the relationships between SCFs (teacher student relations and school discipline) and teachers' job satisfaction and withdrawal cognitions (intentions to quit) are estimated. These SCFs appear to play a protective role with respect to teachers' withdrawal cognitions, and these effects are indirect via their relationship to teachers' job satisfaction.

In paper 3, the relationship between the experience of bullying and students' achievement is addressed. The relationship is found to be indirect, with key roles played by perceptions of school belonging and students' classroom engagement.

Finally, in paper 4 the relationship between civic knowledge and the endorsement of democratic values is estimated. This link is found to be partially mediated by ideological beliefs (authoritarianism), and the role of open classroom discussion (a SCF) as a moderator of these effects is demonstrated.

This work demonstrates that in order to specify theory-driven models of different school outcomes, school climate should be conceptualized as diverse social-contextual effects operating in a complex multivariate setting with mediated and moderated pathways to outcomes.

Table of Contents

List of Figures

List of Tables

Introduction 1

Introduction 2

Paper 1

Paper 2

Paper 3

Paper 4

Paper 4 Appendixes

List of Abbreviations

-2LL	Deviance (-2* Loglikelihood)
AIC	Akaike information criterion
b	Unstandardized estimate
B	Standardized estimate
BIC	Bayesian information criterion
CFI	Comparative Fit Index
df	Degrees of freedom
E	Estimate
e.g.	For example; abbreviation from the Latin *exemplī grātiā*
hOR	Hazard Odds Ratio
i.e	'that is' (to say), from the Latin *id est*
ICC	Intra Class Correlation
LRT	Likelihood Ratio Test
MSEM	Multilevel Structural equation modelling
NTES	National Teacher Evaluation System, conducted in Chile
OECD	Organisation for Economic Co-operation and Development
RMSEA	Root Mean Square Error of Approximation
SE	Standard Error of an estimate
SEM	Structural equation modelling
SES	Socio economic status
SIMCE	Sistema de Medición de la Calidad de la Educación
SRMR	Standardized Root Mean Square Residual
TLI	Tucker-Lewis Index
X^2	Chi Square

Acknowledgements

After long four years, that happened quickly, this journey seems complete. Throughout this time, I’ve collected several memories to be grateful. First of all, I give thanks to my wife Angela. She has bear with my workaholism since the start of the PhD programme till the end of this ride. During the goods, the bad, and stressful ones, her company, have always kept me happy.

For every page I wrote, there was someone on the other side of my screen, who read it, and make sure my work reached a high quality. My supervisor Robin Banerjee, was this person. I’m more than grateful for all of his efforts, to understand what I was trying to say in each my papers and presentations, for helping me to get through all the obstacles during the programme, and support me on every conference I wanted to attend. Thank you Robin. Additionally, I thank all my colleagues at the CRESS Lab, for their comments and reviews, during the progress of this thesis: thanks to Cecilia, Matt, Megan, Helen, Jess, Olaya, Aysenur, and others. Certainly, working in teams makes everything easier.

I also need to thank Rupert Brown and Rod Bond, who both contributed to different stepping stones throughout this journey. When you are a stats geek, I’ve found you feel like the weird kid in the classroom; unless you can share this interest with someone else. Rupert made sure my skills were put at good use during my stay at Sussex, and Rod, have always made me feel welcome, like I was not alone for putting so much attention to statistics and data modelling.

I come from a really small town from Chile, and went to study my undergrad to Santiago, a big city. Moving to Brighton, England, I thought would mimic much of that experience: being alone, and not knowing anyone you could trust for a while. However, Pablo De Tezanos-Pinto, Ximena Poblete and Nicolas Lapport, lend me their best friends respectively. Pablo introduce me to Linda Tipp, who made me feel most welcome in a new University. Xime and Nico, introduce to Farzana Bardai, who show me how to feel more comfortable in a new city. Many thanks to all of you.

It is difficult to write on a topic, that is new in contrast to what you are used to. Thinking about education at large, it is a big challenge. Having colleagues in a similar path contributes to identify your blind spots, open your point of views, and discuss what matters. The whole community at the Society for Education in Chile (SEARCH) at UCL/IOE and the Red de Investigadores chilenos en Educación (RED ICE) provide me with this safe space, to present my work, discuss it and think of it further. Many thanks to you all, in no specific order: Daniel Uribe, Fabian Barrera, Ximena Poblete, Lorena Ortega, Francisca Corbalán, Alvaro González, Ximena Galdames, Maria Jesus Inostroza, Rodrigo Torres, Pablo Torres, Xaviera González, Cota González, Jaime Balladares, Bernardita Muñoz, Maribel González, Mauricio Nercellas, Marcela Ramos, Yolanda Gana, Daniel Leyton, Juan de Dios Oyarzun, Sara Joiko, and the rest of the over 100 members of SEARCH and RED ICE.

Is easier to work and live, when you feel like you have partners in crime at your work place. Wenceslao Unanue and Eleni Ashikali were mine. Their invaluable friendship through the programme, extends over the time and current distance. I am workaholic, yet they made sure I got my fair share of fun. But also, they have been a source of support as a person and as a colleague. Heyla Selim, was my officemate and a true friend, for me and my wife. You made our time in Brighton a pleasant experience. This is also true, for the outside of the office. Michael Leaver, Raul Vallejos and David Martinez, lend me their friendship during these years. Thanks to you all. I always hope to see you all soon. A big thanks also goes to Tomás Saieg, who took the time to go and print this manuscript on my behalf, and take it to the research office. I owe you one.

I also thank to the people from my previous uni, Universidad Católica, who support me in many different ways to get to this stage. I thank David Torres, to push me in the right track, to start working in research. I thank Roberto González, who involve me in so many different research projects, that provide me enough experience to be able to step in complex problems with confidence. I thank Eduardo Barros, who endorse every application I made to get my scholarship. I thank Carlos Cornejo, who taught me how to read between the lines and see the fundamentals. I thank my colleagues in the LAB who have help me in different ways during these years: Daniel Miranda, Pablo De-Tezanos, Hector Carvacho, María Paz Cadena, Héctor Madrid, Siugmin Lay, Daniel Valdenegro, and Patricio Saavedra.

I also thank Yulan Sun and the people at MIDE UC, Flavio Cortez, Mariella Concha, Felipe Zapata, Paulina Morales, Marta Fernandez, Alvaro Zapata, Max Herrada, among others. You all made my stay as a visiting researcher a fruitful experience. I need to thank Jorge Manzi and Daniela Jiménez, from MIDE UC, who secure me a workspace and appropriate working conditions as a post doc, after my doctoral studies, which permit me to prepare for my viva and finish my thesis corrections. Thanks to you all.

Of course, I thank my family. I thank my dad, Juan, who taught the values of being a professional, and my mom, Nirma, who made sure I became one. My sisters, Jimena and Marcela, who always support my departure for my doctoral studies. I thank my parents in law, Celedonio and Victoria, who took me in as one of their sons, gave a me home, and a workspace in Chile. I thank Paula, who welcome me as a brother. I can say, their home, is my favourite workplace throughout this PhD.

I'm the young of two sisters, and never had the experience of having a brother. If I've ever crave such an experience, two of my friends made sure I never need one during my PhD. Alex Torres, support me morally and financially at different stages. Without his support, I would have not attended many of the workshops and conferences I went to. Most my investment in further knowledge used in this thesis, I owe to his support. In spite of the distance and hour difference, Joaquin Reyes, has always been my link to earth. When Angela was still in Chile and I was in the other side of the Atlantic, he was the bridge between us. Thanks to you both for your unconditional friendship.

Finally, this research was possible thanks to the financial support of the Programme of Advance Human Capital, National Commission of Scientific and Technological Research of Chile (CONICYT) PAI/INDUSTRIA 72111272, who sponsored these studies, with a full scholarship.

Statement of authorship and contributions to the work

I mainly use the pronoun "I" in the first few sections and at the end of the thesis, whereas in the empirical papers, the pronoun "we" is used to refer to the authors. All papers within this thesis are co-authored with Robin Banerjee, my supervisor. However, in spite of this explicit collaboration, the main body of work is attributed to myself. To clarify the joint use of "I" and "we" throughout this text, it should be noted: the full design of the research work, literature review, statistical analysis, and writing of chapters and papers were my responsibility. Robin Banerjee discussed my choices and made comments on the writing, as part of the supervision process. Yet, the present work's authorship, as a doctoral thesis, should be attributed to me.

Chapter 1

Introduction to research on school climate: Historical and conceptual overview

The overall aim of this programme of doctoral research was to evaluate how and to what extent school climate is related to key school outcomes such as student achievement and teacher job satisfaction and retention. Many may agree that a positive school climate is central or critical for school success, but school climate is an elusive concept with more than one meaning, and several different ways to measure it. In order to assess the relationship between school climate and school outcomes, we must first resolve conceptual and practical issues. There is a need to clarify the nature of school climate, and consider how best to assess its relation to school outcomes. In this chapter, we review the basic roots of 'school effectiveness' research, a key strand of academic enquiry which seems to have led to school climate being seen as critical for school success. Then, we will outline the conceptual history of school climate within other research strands, before arriving at a conceptualisation of school climate that will guide the present empirical research. In the next chapter, we review the problematic issues commonly faced within school climate research, and we delineate alternative strategies for dealing with these, in order to answer different research questions.

Variations in school results

Why do schools have different results? What makes one school more successful than another? These are common questions within educational research and educational policy, and are relevant questions for parents and children when choosing schools, or indeed for teachers who are thinking about possible workplaces. Different factors constitute possible alternatives for answers, such as school intake and teaching quality, among many others. However, the disentanglement of the different factors involved in accounting for school outcomes is not an easy task.

Different disciplines have been involved in answering these questions, including economics, education, sociology, and psychology. A more specific disciplinary framework devoted to these questions is 'school effectiveness' research. Historically, the main aim of this research endeavour has been to identify the key factors that explain school outcomes. Specifically, it: a) addresses performance differences between and within schools, b) aims to identify the malleable factors that enhance school results, and c) frequently uses student achievement scores as the key criterion measure of school performance (Luyten, Visscher, & Witziers, 2005). Understanding the starting point of this discipline is relevant to answering the questions of what factors account for educational outcomes, and their relation to the concept of school climate.

Perhaps a more fundamental question is how much schools can ‘make a difference’ to student outcomes, especially when comparing children from different socioeconomic and family backgrounds. The works of Coleman and colleagues (1966) and Jencks (1975) are considered a crucial point in history for educational research in this regard. Coleman and colleagues (1966) concluded in their report that only a small portion of student achievement could be attributed to school factors, in contrast to family background. The work of Jencks (1975) reinforced this conclusion, in which a strong link was found between socioeconomic status (SES) and student achievement from one generation to the other. The author thus argued that it is not advisable to rely on education as a ‘means to an end’, in order to achieve social wealth distribution (S. M. Miller, 1973). These conclusions gave birth to the ‘schools *do not* make a difference’ interpretation of their work, to which school effectiveness research largely reacted (Luyten et al., 2005; Wyatt, 1996).

Rutter and Maughan (2002) asserted that ‘can schools make a difference?’ was not the appropriate question, at least not as it had been interpreted at that time. The burden of proof had been placed at the level of population variance. That is, the expectation was – and sometimes still is – that schooling would reduce the differences between socioeconomically advantaged and disadvantaged pupils, at the level of the population variance. They proposed to change the frame of reference for this question, and ask if the amount and quality of schooling children receive influences their progress during childhood and adolescence. This changes our perspective on judging differences between schools. The works of Edmonds (1979a, 1979b, 1982) on school improvement, and extensive studies of school climate from Brookover (1979) in the US; and the works of Reynolds (1982) and Rutter and colleagues (1982) in the UK took this perspective. These later approaches placed much more attention onto the context and the ‘learning environments’ of schools in order to account for school results.

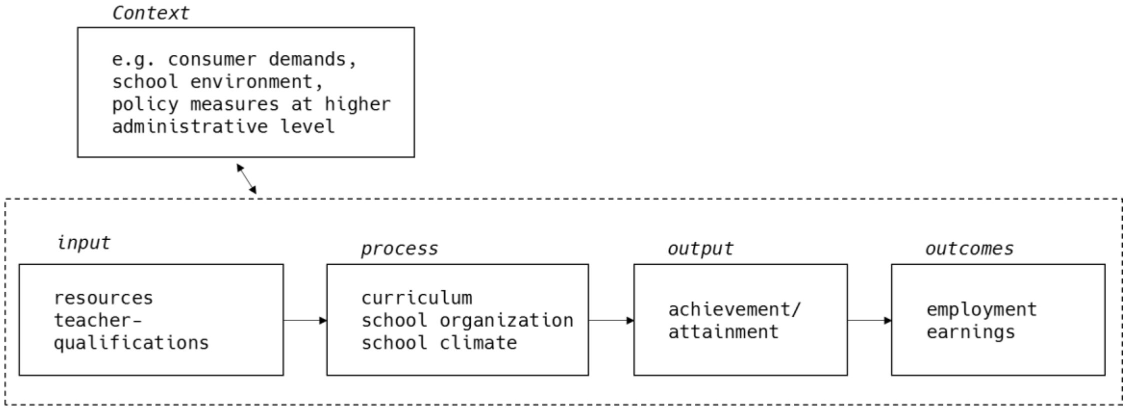
School factors and students' academic achievement: An input-process-outcome model

School effectiveness research has drawn attention to a range of factors such as professional leadership, shared vision and goals within the school community, characteristics of the learning environment, quality of teaching, positive reinforcement, monitoring of student progress, and pupil rights and responsibilities (Reynolds, 1997).

Scheerens, Glas, & Thomas (2003) propose a summary of 14 effectiveness enhancing factors, including: high teacher expectations, educational leadership, cohesion among the staff, curriculum quality/opportunity to learn, school climate, evaluative potential, parental involvement, classroom climate, effective learning (time use), structured instruction, independent learning, differentiation, keeping records on students' progress, feedback and reinforcement. Integrated models of school effectiveness, such as the dynamic model of school effectiveness (Creemers & Kyriakides, 2007, 2010; Kyriakides, Creemers, Antoniou, & Demetriou, 2010), focus on similar factors and include additional factors such as teachers' and students' values, alongside higher-level factors such as school policy, and national and regional levels of educational policy.

Many of these factors can be ordered in terms of an 'input – process – outcome' model for schools. For example, school intake (e.g., students' family background) can be interpreted as an input for the schools. Teacher qualifications can also be considered an input for this model. A typical outcome for the schooling process might be general literacy, or reading ability. The intervening process, then, would be all the school characteristics and practices that might have some bearing on the outcome variable. These process factors may account for the relationship between input and output (mediation) or they might interact with the input (moderation). Yet, all of this is an oversimplification. As I have mentioned previously, the task of disentangling the school effects from the family and school SES factors is not an easy task, and other contextual factors should not be overlooked. An enriched version of the input, process and outcome model is presented by (Scheerens, 1990) and shown in Figure i1.1

Figure i1.1: Context - input - process - output - outcome model of schooling



Process indicators have the role of offering plausible explanations of why schools, or schools systems, differ in their results (Scheerens, 1990). Process indicators are ways to characterize what happens in the black box of the school process. In this sense, most of the previously mentioned school effectiveness factors are process indicators. However, there is some conceptual fluidity in the consideration of inputs and process variables (Reynolds, 1997). For example, teaching quality could be operationalized as teacher qualifications, hence an input factor, but it could alternatively be portrayed as a process factor if it is operationalized in terms of observed instructional practices.

The wide range of school factors that have been considered can themselves be located on multiple levels. The dynamic model of school effectiveness (Kyriakides et al., 2010) distinguishes multiple nested levels. The first is the teacher-student relation, which happens in the classroom; this level constitutes the learning situation. However, this relation is nested within schools, and consequently schools are nested in wider contexts. The context level permits researchers to compare different areas, regions, and nations. The whole idea of this model is to account for nationwide policies, which influence school policies, which in turn should guide teaching practices with pupils. Of course, these different factors have different degrees of proximity to the learning situation, and it can be reasonably expected that classroom level factors have more significance than the school and system level factors to explain educational outcomes (Kyriakides & Creemers, 2008; Scheerens, Vermeulen, & Pelgrum, 1989).

By using the list of 14 effectiveness factors from Scheerens and colleagues (2003), we can make a distinction (see Table i1.1 and Table i1.2).

Table i1.1 Fourteen school effectiveness-enhancing factors Scheerens et al. (2003)

	Effectiveness Enhancing Factors	Description	School	Classroom	Aspect
1.	achievement orientation/high expectations/teacher expectations	Teachers hold high expectations of their pupils' educational attainment.	X	X	SEC
2.	educational leadership	Head teachers and principals lead school practices by monitoring progress, counselling teaching practices, establishing clear objectives, prioritising pedagogical work over bureaucratic tasks, and stimulating cooperative relationship between staff to achieve school mission.	X		SEC
3.	consensus and cohesion among staff	Teaching staff cooperate with each other and establish consistent practices across grades and subjects, on evaluative practices, curriculum, and teaching practices.	X		SEC
4.	curriculum quality/opportunity to learn	What is taught in each subject. This could be assessed by considering the list of content included in an educational year; considering also time allocations with subject variety.		X	IC
5.	school climate	School pride, student satisfaction, orderly learning environment and achievement oriented values.	X		SEC
6.	evaluative potential	Use of evaluation of progress of students learning and school overall function, to inform school management and teaching practices.	X	X	IC
7.	parental involvement	School has regular contact with parents. They can drop in, and parents have a supporting attitude to the schooling.	X	X	SEC

Note: SEC = socio-emotional component, IC = instructional component

Table i1.2 Fourteen school effectiveness-enhancing factors Scheerens et al. (2003) (continuation…)

	Effectiveness Enhancing Factors	Description	School	Classroom	Aspect
8.	classroom climate	Positive relationships within the classroom (teacher-student, student-student), order and work attitude in the classroom, and satisfaction		X	SEC
9.	effective learning time (classroom management)	Exposure time to educational treatment at school		X	IC
10.	structured instruction	Structured and well prepared lessons. Subject matters are divided in units, which are congruently sequenced for learning. It responds to a model of 'good teaching'.		X	IC
11.	independent learning	How much time and opportunities students have to realize their own chosen work for a subject matter.		X	IC
12.	differentiation, adaptive instruction	Adaptive instruction for students with different abilities.		X	IC
13.	keeping records on pupils' progress	How formalize is the school system on student's records.	X		IC
14.	feedback and reinforcement	Rapport to pupils in connection with their achievement. How feedback is conducted in the school.		X	IC

Note: SEC = socio-emotional component, IC = instructional component

There is a contrast between those factors that manifest at the school level and those that relate more to classroom level. Additionally, each indicator can be roughly classified into those which are closer to teacher practices, and those which present a more indirect nature. Chávez (1984) makes a distinction between indicators within the research of classroom climates: between low and high inference indicators. The ‘low inference’ indicators are direct measures and seem less controversial. For example, classroom noise, teacher time use, and classroom teacher/student ratio are all low inference measures. In contrast, in the case of ‘high inference’ indicators, the subjective nature of what is measured is at play. In this regard, students’ sense of belonging to the school is a high inference measure. In the latter case, the subjective experience of the school members is the source of variance, though it still remains attributable at least in part to the school context.

Rubie-Davies (2014) makes a related distinction, when the authors argue that classroom climate is the result of two components: the instructional and the socio-emotional. The first represents the results of teachers’ pedagogical decisions, whereas the latter is the result of the interaction between teacher and students, and the interaction of students and their peers. Most of the time, the socio-emotional component of school climate is assessed or captured by high inference measures. In Table i1.1 and Table i1.2 below, factors are marked with an ‘X’ under columns ‘School’ and ‘Classroom’ to point out whether these could be observed at the school and/or classroom level. Additionally, in the last column of the table, factors were classified as socio-emotional components (SEC) or instructional components (IC).

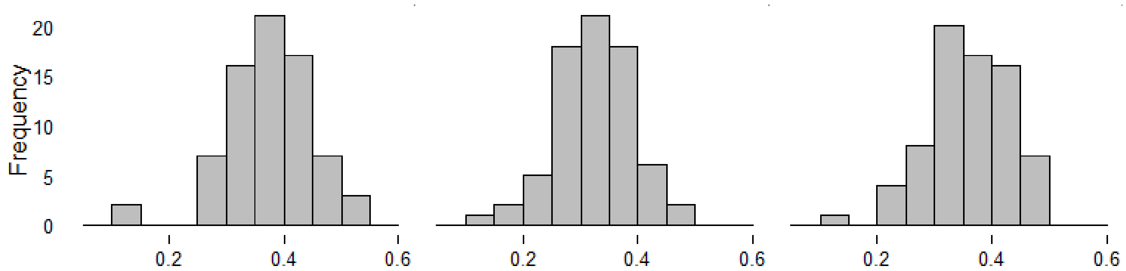
The role of socioeconomic status

As a quick scan of the many national and international ‘league tables’ can attest, attainment results vary across schools. What factors account for this differences? Multiple factors could be considered: wealthy areas versus disadvantaged areas, school socioeconomic composition, proportion of immigrants in the school intake whose native language is different from their schoolmates, public versus private schools, single-sex versus mixed schools, proportion of certified teachers, number of hours in effective teaching and so on. Out of all these factors, perhaps the most documented factor linked to school differences is socioeconomic status (SES). Yet even here, the picture is not quite as straightforward as it might appear. The between-school association of SES and academic achievement is very substantial, with a meta-analytic correlation estimate of about .73, but at the level of the individual, estimates have shown a much weaker effect of .22 (White, 1982). More recent meta-analyses have shown similar results. Sirin (2005) showed correlations of .67 and .28, for the school and the individual level data respectively.

Estimates of correlation of the school mean SES and school results are a blended or conflated estimate of the school and students inference level. School compositional effects or contextual effects (van Ewijk & Sleegers, 2010) estimate the specific effect of school SES composition while controlling for the students' SES background. Meta-analytic estimates of this parameter are of r=.32 (van Ewijk & Sleegers, 2010). This parameter is a measure of the expected difference in student achievement between two students with a similar SES background who attend schools with different school mean SES (Caro & Lenkeit, 2012; Willms, 2010).

As with any other meta-analytic estimates of this nature, these estimates are both moderated by how SES and academic achievement measures were taken, a caution shared by all the authors (White, 1982; Sirin, 2005; van Ewijk & Sleegers, 2010). A way to encompass this limitation and get a general picture of the relation between these two factors is to estimate these effects across different countries at one-time point with the same instruments. Such an attempt, is displayed in the following figure.

Figure i1.2 Histogram of country effect sizes correlations between socioeconomic background and student's achievement (PISA, 2009)



Math Scores Reading Scores Science Scores

Note: Author's estimates.

Secondary analysis from the Program for International Student Assessment (PISA) from 2009 (OECD, 2010a) permits this exercise. Correlations between socioeconomic background of pupils and their achievement yield average effects of .31

to .37 of socioeconomic status (SES) on reading, math and science scores, for random samples of 15-year-old students in various countries. Figure i1.2 shows that even though the average of the estimated effects of all the random samples per country is within the range of the meta-analytic estimates, there are some countries in which the relation of these factors is lower. Nevertheless, more than 60% of the countries have effects over .3 for the three academic subjects (math, reading, and science).

In general, most of the school effectiveness research evaluates effects *after* controlling for SES. The basic idea is to identify factors that influence achievement and other educational outcomes *over and above SES effects*. Scheerens & Bosker (1997) conducted a meta-analysis of over 150 studies of school effectiveness including contexts from Europe, North America, Australia and third world countries, including primary and secondary schools. They found that school factors accounted for 19% of the attainment variance, prior to controls for intake characteristics (socioeconomic factors). When the latter were included, the observed net-effect of school factors was moderated, accounting for only 8% of the variance in school attainment.

Comparison between countries, using data from Trends in Mathematical and Science Study from 2003 (Mullis, Martin, Gonzalez, & Chrostowski, 2004), which targets random samples of students at grade 4 (10 to 11 years old) regarding their achievement in Maths and Science in 25 countries, also shows that in most cases family background explains more variance than schools do (Chudgar & Luschei, 2009). The authors, using an ordinary least square (OLS) regression approach on Math scores at the country level, found that adjusted ΔR^2 was larger for family effects than for school effects in 18 out of the 25 countries. Nevertheless, these effects vary between richer and poorer countries, especially for countries with higher inequality. Specifically, the impact of school decreases (relative to families) as national income increases. Complementing this, there is a trend in which the higher the income inequality of the countries, the larger the effects of the schools.

It is important to recognise that school factors may also *interact* with SES, such that effects of socioeconomic disadvantage on achievement may be moderated by school factors. For example, Zhao, Valcke, Desoete, & Verhaeghe (2012) showed that low income students achieve higher results in schools with higher family socioeconomic status in contrast to students who are enrolled in schools with lower and average family socioeconomic status. These effects could involve a variety of practices located at the level of the school. For example, Lee and Bryk (1988) compared the effects of tracking practices between Catholic and public schools and found that Catholic schools show a greater tendency to assign their students to academic tracks (in contrast to remedial courses, and vocational training). When these effects were compared with socioeconomic background variables, researchers found that in Catholic schools' track placement is less influenced by family background, in comparison to public schools. Hence, track placement can potentially play an important moderating role in the link between students' background and academic achievement.

Much of the attention school effects receive for policymaking is because school compositions are partially driven by system level properties, such as school tracking, school selectivity, school choice and urban segregation. All of these have potentially different effects on students' academic outcomes, conditioned by their socioeconomic status (Elacqua, 2012; McEwan, Urquiola, & Vegas, 2008; Valenzuela, Bellei, & Ríos, 2014). Adding in a more fine-grained analysis of subjective variables relating to the school context can enhance our explanation of variance in school outcomes even further. For example, Gutman and Midgley (2000) showed that an interaction between students' sense of belonging to school and parent involvement with school predicted variance in academic achievement for disadvantaged students. Schools exhibiting a higher sense of belonging among students and more parental involvement tended to have higher results than those schools that lacked these features.

In summary the cited research supports the idea that school practices do influences school achievement, and also that social aspects of schools – and stakeholders' subjective experience of those dimensions – interact with students' background characteristics to influence academic achievement.

Problems with school climate within school effectiveness research

Many criticisms have been made of school effectiveness research, neatly summarised by Goldstein and Woodhouse (2000). The most common criticism of school effectiveness research has been its lack of theory (Lauder, Jamieson, & Wikeley, 1998; Thrupp, 2001; Goldstein & Woodhouse, 2000; Luyten et al., 2005). Scheerens (2013) compiled 109 international studies of school effectiveness, and only six of these articles were classified by the author as theory driven. Correspondingly, school effectiveness research has been accused of fishing for correlations among arrays of indicators (Coe & Fitz-Gibbon, 1998); once some of these show significance, they are later on reified into broader concepts such as ‘leadership’ and ‘ethos’ by piling up different items using common sense meaning (Lauder et al., 1998).

School effectiveness is appealing to public policy, thanks to the reengineering assumption: once effective enhancing factors are identified from the successful schools, those practices could be imported and replicated in less successful schools in order to improve, or to reengineer its processes (Lauder et al., 1998). It is not hard to see why school climate might become the all-encompassing, fundamental basis for this recipe (Sandoval-Hernandez, 2008).

This is the crucial problem of research on school climate from within school effectiveness research: it is characterized, by most critics, as just a collection of items which differ largely from one study to another (S. I. Miller & Fredericks, 1990). These constitute no more than a pure false entity, a reification, a category mistake in the purely Rylean sense (Ryle, 1949), according to its critics. Of course, reifications and vague concepts are not a unique issue of school climate research; they are a common problem within the social sciences in general (Billig, 2013). Nevertheless, they pose a major challenge for the present research into school climate.

The present research aims to address some of these challenges, by providing theory driven models which assess the relationship between school climate and school outcomes. Yet, before we can do this, some conceptual distinctions need to be made in order to offer a stance in which school climate dimensions can be operationalized within a conceptual model. In the next section, we review existing school climate approaches.

A history of approaches to school climate

School climate is a loose concept. Most of its definitions tap into different construct and concepts including shared perceptions, feelings, beliefs and attitudes, among others. A very wide array of characteristics has appeared under the school climate banner. In fact, school climate surveys are used to capture aspects of teacher efficacy, principal leadership, school safety, student activities (Mo, Yang, & Hu, 2011), student peer relations, ambient noise, overall satisfaction, and student-teacher relations, to name a few (Freiberg, 1998, 1999). All in all, school climate has been used as way of describing the school experience (MacAulay, 1990) and the school culture (Lindahl, 2006) in the broadest sense.

The first common metaphors which underpinned the notion of social climate concerned the ‘personality’ of an organization (Halpin & Croft, 1963) and its organisational ‘health’ (Hoy & Fedman, 1987). These early approaches to school climate were similar to those approaches to organisational climate more generally (i.e. group climates in workplaces), inheriting instruments, theory and methods (C. S. Anderson, 1982). Similar to the organisational literature, school climate research has developed a multidimensional approach to assess different factors related to different school outcomes. Where industrial psychology was interested in understanding commitment, satisfaction with the workplace, and performance (Carr, Schmidt, Ford, & DeShon, 2003), school climate has focused on students’ and teachers’ engagement and satisfaction with the school and its performance (Ostroff, 1992). In a general sense, both notions of organizational or social climate share the epistemic model in which a person is embedded within a group, nested within an organisation, in a particular context. Thus, the experience of members of the same school (i.e., organization) shared the object of the experience.

Over 100 years have passed since Perry (1908) introduced the notion of school climate. This author was a school principal who wrote one of the earliest books on the topic: *The Management of a City School*. In this work, Perry described the many roles and concerns a school principal has, and school climate was one of them. In an indirect way, Perry suggested to teachers and principals the need to be aware of the importance of the setting, or the school *atmosphere* (Cohen, McCabe, Michelli, & Pickeral, 2009; Zullig, Koopman, Patton, & Ubbes, 2010; Zullig, Koopman, & Huebner, 2009; Freiberg, 1999). In Perry’s view the proper school atmosphere was established as a moral ethos, an expression of the school values. The author targeted this notion as a key factor for school management, which involved the principal, the teachers and the students. Perry viewed this group atmosphere as part of a tradition. As such, it was not built in a day; it was the result of habitus and group morality, reinforced by senior pupils to younger pupils, expressed between students from different schools, and influenced by group activities and participation such as the school sports team, students’ association (i.e., student union), student clubs (e.g., music, literary, scientific), among other school social practices. He referred to a ‘school spirit`, involving aspects that schools develop as a community.

The importance of an *atmosphere* in understanding human behaviour was also highlighted by the seminal works of Lewin and colleagues (K. Lewin, Heider, &

Heider, 1936), which emphasise the relation between the person and its environment, between the individual and its group. Murray's model of needs and environment suggests that the environment 'presses' the individual to frustration or to satisfaction (Murray, 1938). Both accounts highlight that individuals' behaviour in group settings is rarely accounted for just by individual differences. People moderate their behaviour in relation to the groups of which they are a part of (Fraser, 1989), and people tend to perceive aspects of their everyday experience in collectives. The school scenario and the work organisation seem to be no exemption from this phenomenon. Nevertheless, the proper conceptualisation of this social aspect within school settings seems progressive: from the factual account of group composition to the inclusion of rather less tangible group values and attitudes.

The notion of environment in this kind of inquiry was not formalized until the 1950s, with the work of March and Simon (1958) and Argyris (1958). This work covered diverse aspects of human behaviour in institutional settings, using notions such as morale and attitudes to account for productivity and turnover in organisations (Zullig et al., 2009). Later on, the works of Halpin & Croft (1963), Walberg & Anderson (1968), Moos (1979) and Brookover and colleagues (1978) would represent the foundational applications of school and class environment in this sense.

During the 1960s and early 1970s there was a special emphasis on the study of schools, including school's intake composition to account for school success. Coleman and colleagues' report (Coleman et al., 1966), mentioned earlier, is considered foundational in this trend. In this early work, students' ethnicity and socioeconomic differences were often considered as proxies for school climate, conforming to models of input-output theories (C. S. Anderson, 1982). However, other researchers were beginning to think about school climate in terms of organisational features. Halpin and Croft (1963) for example created the Organizational Climate Description Questionnaire (OCDQ), constituted by 64 items. These items were developed from an initial bank of one thousand items, which were refined by content analysis and factor analytic methods in pilot studies, avoiding redundancy and assuring thematic coverage. The 64 items were organized into 8 subscales to characterise teacher and principal behaviours. The following table displays examples of subscales for each domain.

Table i1.3 Organizational Climate Description Questionnaire (OCDQ) domains

Dimension	**Description**
Teachers Behaviours	
a) Disengagement	Teachers talk about transferring from this school
b) Hindrance	Amount of routine non-pedagogical task (‘Teachers have too many extracurricular requirements’)
c) Esprit	Teacher’s morale, concerning the degree to which their social needs are being satisfied along with a sense of accomplishment in their job (‘In staff meetings there is a feeling of “let’s get things done”
d) Intimacy	Teachers’ enjoyment of friendly social relations (‘Teachers talk about their personal life to other staff members’).
Principal Behaviours	
a) Aloofness,	How much the relationship was formal and impersonal (‘Teachers are contacted by the principal every day’);
b) Production	Emphasis, referring to the inclination of the principals to be highly directive and task oriented, and characterised by close supervision (‘The principal ensures that teachers work at their full capacity’)
c) Thrust,	Task oriented inclination which is perceived positively by teachers, whereby the principal leads by example and motivates teachers (‘The principal tells teachers of new ideas he has run across’)
d) Consideration,	referring to how much principals show personal consideration with teachers (‘The principal tries to get better conditions for teacher’).

These different dimensions were proposed to describe the personality of the school organisation. In contrast, the work of Brookover and colleagues (1978) emphasised the beliefs and perceptions of norms within the schools held by the different members of the school community: students, teachers, and school principals. All of these stakeholders responded to items related to educational expectations, quality of education, academic norms, and academic futility. Academic futility was a novel concept introduced by Brookover and colleagues, directly inspired by the Coleman Report (Coleman et al., 1966) and its construct of ‘sense of control’. But it was directly designed to assess the level of the student sense of control within the school (Houtte & Stevens, 2008). Academic futility expresses a feeling of lack of control over academic results (item sample: ‘People like me will never do well in school, even if we try hard’). According to Brookover et al. (1978), this factor showed a correlation of .77 with school mean achievement for the random sample of Michigan state schools, and effects of .69 for schools with majority black students, in contrast to .51 for schools with majority white students.

These two approaches express an important divergence within school climate research. In this field, one strand of studies focuses on staff organisation, whereas others focus on shared feelings, values and norms of the whole school (Griffith, 1997). This difference also echoes the distinction between those studies that only included the relationship between the school principal and the teachers, and the studies that included additionally students' responses. The first trend of research echoed the methods and approaches of organisational research, while the latter is the consequence of sociological approaches, in which the school success (output), was a function of a social process in which the school is a cultural system of social relationships among the school community (family, teachers, students and peers) (C. S. Anderson, 1982).

According to C. S Anderson (1982), a complementary approach emerged between the 1960s and 1970s: the ecological approach. Similar to sociological approaches, it includes process indicators, but adds wider issues relating to the distribution and allocation of resources, and the temporal and physical aspects of the environment. It attempts to account for the system as a whole, and views all variables as potentially modifiable for the benefit of the student outcomes. Advocates of the ecological approach are Barker (Barker & Gump, 1964; Barker, 1963), Goodlad (Goodlad & Sava, 1975), Eggleston (1977) and Moos (1979). For example, Eggleston (1977) focuses his attention on high-level comparisons of higher education enrolments between different local authorities in England and Wales, which were closely related to school resources allocation from school provisions within these local authorities. These broader accounts are complementary to the process indicators that the sociological approach undertook. However, these approaches put attention on possible moderating factors between larger school areas.

In the 1990s an array of studies started to make the distinction between the school and class environment (Zullig et al., 2010). Griffith (1995) discussed what is the most appropriate level of inference for school climate: the school as a whole or the classroom. The author argued that for the case of schools where lessons are held in different rooms with different teachers, in which students move from room to room, it makes sense to posit the school as the unit of analysis. On the other hand, for schools in which a large proportion of the teaching is held by a single teacher in a single room with a constant group of pupils, then it would be natural to identify the unit of analysis for climate at the classroom level. This reiterates the distinction we made earlier when considering the dynamic model of educational effectiveness (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2008), in which special attention is placed upon the learning situation in particular, as the most important level of analysis to account for academic achievement.

School climate within classrooms

Other authors have traced back to earlier years the research into school social climates in terms of classroom-level processes. Chávez (1984) undertook a review including references from the late 1920s to the 1960s. In his article, he targets Dorothy Thomas (Thomas, 1929) as the earliest researcher to observe class environments. This earliest work set up the standard for the ensuing development of classroom environment research: objective records of specific social/test situations and accurate observation of children's behaviour. Later on, the work of social psychologists showed the relationship of leadership and group social climate to aggressive behaviour and productivity of groups (Lippitt & White, 1943; Kurt Lewin, Lippitt, & White, 1939; Kurt Lewin & Lippitt, 1938). Among 10- to 11-year-old children working in clubs, democratic groups had moderate productivity outcomes, regardless of the presence of the leader, whereas authoritarian groups dropped their rate of productivity by 45% approximately, conditioned by the presence of their leader (Lippitt & White, 1943) .

Throughout this investigation of classroom-level processes, social interactions have been seen as critical. Anderson and colleagues (H. H. Anderson & Brewer, 1945, 1946; H. H. Anderson, Brewer, & Reed, 1946) studied the effects of teachers' personality in the classroom and its effects on students' behaviour. The 1950s, according to Chávez (1984), were more theoretically driven, basing their models on the main tenets of Murray's (1938) need-press model. Thelen's (1950) educational model relied on the analysis of pictures and recordings to measure classroom environments. Withall (1949) was the first author to mention directly the concept of 'climate'. He described the learning situation, in which teacher and students interact, as characterized by a particular *social emotional climate:* climate is considered to represent the *emotional tone* of face-to-face interactions. This is related to the degree of acceptance expressed by members of a group regarding each other's needs and/or goals (Withall, 1949).

Applied research from those years had explicit hypotheses involving conditions of this social dimension. For example, Bovard (1951) hypothesised that classroom interaction would influence classroom group cohesion. Among college students, classrooms richer in students' verbal interaction should promote cohesion among different students of the class group, overcoming differences in terms of race and socioeconomic characteristics. They found result in favour of this assumption. Along similar lines, the work of Buswell (1953) studied the relationship between acceptance and rejection among peers on the one hand and academic success on the other. It was assumed that acceptance levels of children would be positively related to academic achievement. In a group of 286 kindergarten children and 321 fifth grade pupils, she found supporting evidence for this link, only for later grades but not for kindergarten children.

This emphasis on social interaction as one of the key 'process indicators' related to school climate was complemented by the emergence of school climate surveys referred to earlier. These instruments were 'high inference' in nature (Rosenshine & Furst, 1971). That is, instead of relying on direct and 'objective' observation of events made by raters or judges (checklists, rubrics, frequency counts, true or false marks of events), these were measures of appraisals, index scores of agreements with items describing classroom events or characteristics, typically measured on Likert-type scales (Clason, Dormody, & Scales, 1993).

In fact, the preference for the high-inference approach to measurement was to some extent empirically based. Apart from being a more resource-efficient measurement approach, the subjective and affective dimensions of school climate turned out to be more predictive of school outcomes than the objective record of activities and environmental features. For example, the Observation Schedule and Record tool (OScAR) (Medley & Mitzel, 1958, 1963) was developed as a template to observe classroom practices but also included measures of socioemotional climate. Further research with this instrument showed that the emotional climate scale was more sensitive and more related to students' and teachers' behaviours than the other scales in the measurement system. This would a set a precedent in classroom research, in which the emotional tone of classroom climate was considered a crucial factor (Chávez, 1984).

A review of school climate domains

Freiberg (1998) used the expression 'let me count the ways' for the many existing operationalisations of school climate, at the school and classroom level. He suggests no single factor of school climate exhaustively captures social climate as a concept. As such, it is difficult to provide a single definition that is inductively satisfactory for the full array of climate instruments. Cohen and colleagues (Cohen et al., 2009) prefers to stick to the notion of quality and character of school life. This definition is broad enough to include in the same compendium constructs such as 'cohesiveness within classrooms' and constructs such as 'academic futility'. There is no fixed list of constructs to identify definitively what is school climate, and what is not school climate (Cohen et al., 2009; Cohen, 2012; Freiberg, 1999; Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013).

Table i1.4 Summary table of classroom environment scales (from Fraser, 1989)

			Dimenions according to Moos' scheme		
Insrtument	Level	Items per scale	Relationships	Personal developtment	Maintenance and change system
Learning Enviroment Inventory (LEI)	Secondary	7	Cohesiveness Friction Favouritism Cliqueness Satisfaction Apathy	Speed Dificullty Competetiveness	Diversity Formality Material enviroment Goal Direction Disorganization Democracy
Classroom Enviroment Scale (CES)	Secondary	10	Involvement Affiliation Teacher Support	Task orientation Competition	Order and organization Rule clarity Teacher control Innovation
Individualized Classroom	Secondary	10	Personalization Participation	Independence Investigation	Diferentiation
My Class Inventory (MCI)	Primary Lower Secondary	6-9	Cohesiveness Friction Satisfaction		
College and University Classroom Enviroment	Higher Educati	7	Personalization Involvement Cohesiveness Satisfaction	Task orientation	Innovation Individualization
Science Laboratory Enviroment	Senior seconda Higher Education	7	Teacher support Involvement Cohesiveness	Open endedness Integration	Organization Rule clarity Material enviroment

For the classroom level, Fraser (1989) reviews 6 different high inference instruments of classroom emotional climates, which were developed from the 1960s through to the late 1980s. The scheme set out by Moos (1980) helps us see that most of the instruments account for: relationships of students and teachers, in terms of both quality and involvement; personal growth and self enhancement; and changes to or maintenance of the classroom environment, including aspects of rules, organisation, and formality (see Table i1.4).

Meanwhile, as shown in Table i1.5 below, Cohen and colleagues (Cohen et al., 2009) propose that school climate can be understood in terms of four spheres or thematic domains: safety, teaching and learning experience, relationships among school members, and environmental and structural features of the school.

Table i1.5 The four school climate domains of (Cohen et al., 2009)

Domain	*Dimensions*
Safety	**a. Physical** (e.g., crisis plan; clearly communicated rules; clear and consistent violation response; people in the school feel physically safe; attitudes about violence) **b. Social-emotional** (e.g., attitudes about individual differences; students' and adults' attitudes about and responses to bullying; conflict resolution taught in school; belief in school rules)
Teaching and Learning	**a. Quality of instruction** (e.g., high expectations for student achievement; all learning styles honoured; help provided when needed; learning linked to "real life"; engaging materials; use of praise/reward; opportunities for participation; varied teaching methods; instructional leadership; creativity valued) **b. Social, emotional and ethical learning** (e.g., social-emotional and academic learning valued/taught; varied "intelligences" appreciated; connections across disciplines) **c. Professional development** (e.g., standards and measures used to support learning and continual improvement; professional development is systematic and on-going; data-driven decision making linked to learning; school systems evaluated; teachers feel that this is relevant and helpful) **d. Leadership** (compelling and clearly communicated vision; administrative accessibility and support; school leaders honour people at school)

Table i1.6 The four school climate domains of (Cohen et al., 2009) (continuation…)

Domain	*Dimensions*
Relationships	**a. Respect for diversity** (positive adult-adult relationships between/among teachers, administrators, and staff; positive adult-student relationships; positive student-student relationships; shared decision-making; common academic planning opportunities; diversity valued; student participation in learning and discipline; peer norms linked to learning, cooperative learning, conflict-violence prevention; being able to say "no") **b. School community & collaboration** (mutual support and on-going communication; school-community involvement; parent participation in school decision-making; shared parent-teacher norms vis-à-vis learning and behaviour; student family assistance programs) **c. Morale and "connectedness"** (students are engaged learners; staff are enthusiastic about their work; students connected to one or more adults; students/staff feel good about school and school community)
Environmental -Structural	(cleanliness; adequate space and materials; inviting aesthetic quality and size of school; curricular and extracurricular offerings)

Cohen and colleagues' (2009) classification of school climate factors relies heavily on the application of school climate research to public educational policy, and connects with the different research traditions described earlier. For example, the safety dimension is closely related to school programmes concerning violence and bullying, in which school climate plays a critical role. Here the notion of school climate in use is closer to Brookover's (1979) approach, with a primary focus on the attitudes, beliefs, values and norms within the school (McEvoy & Welker, 2000).

In contrast, the domain of teaching and learning is closer to the research tradition of classroom environment research. This includes the assessment of teacher practices through student views, the emotional tone of the classroom (Withall, 1949) and teachers' and students' expectations (Brookover et al., 1978). All of these components are thought to interplay in the learning situation.

On the other hand, school relationships, in a general sense, are closer to the focus on the quality of the social interactions among the members of a school present in the work of Bryk and colleagues (Bryk & Schneider, 2004; Bryk, Sebring, & Allensworth, 2010) and Hoy and colleagues (Forsyth, Adams, & Hoy, 2011; Hoy, 2012). These two research groups largely converge on the main argument that quality of interpersonal relationships are crucial for school improvement. Namely, school trust among school members is thought to play a critical role to explain differences between schools, implementation of school reforms, school effectiveness, and school improvement.

Finally, environmental and structural appraisals of schools are much more related to the ecological approach, tackling questions regarding the relationship between school resources and student outcomes, and factors such as size, distribution of the space and quality of school environmental provision (Thapa et al., 2013).

Table i1.7 Historically common school climate domains, according to Zullig, et al (2010)

Domain	*Domain Variations*
Order, Safety, & Discipline	Perceived safety Respect for peers and authority Knowledge and fairness of disciplinary policies Presence of gangs Accomplishment and recognition
Academic Outcomes	Sense of academic futility Academic norms Academic instruction Overall satisfaction with classes Future and present evaluations of performance
Social Relationships	Teacher-student relationships Interpersonal relationships Student-peer relationships Helpfulness of school staff
School Facilities	School temperature Classroom arrangement Ambient noise School, classroom, and grounds condition School decorations
School Connectedness	Excited, enthusiastic, and engaged learners Feelings about school Students feel valued for their input

As shown in Table i1.7 below, Zullig and colleagues (Zullig et al., 2010) reach a similar summary. In contrast to previous summaries, these authors include a collection of different constructs within each domain.

The most recent school climate review, conducted by Wang and Degol (2015) proposes a very similar classification. The strength of this last categorization of school climate domains lies in their research strategy. The authors used a systematic review approach, collected over 327 references and used the 50 most cited articles as a guide to build up this classification. Additionally, these authors consulted with scholars within school climate research to rectify the categorization for underrepresented areas.

Table i1.8 School climate domains and dimensions, according to Wang and Degol (2015)

Domain	*Dimension*
Safety	Socio Emotional Discipline and Order Physical
Community	Partnership Quality of Relationships Connectedness Respect for diversity
Academic	Leadership Teaching and learning Professional development
Institutional environment	Environmental Structural Organizational Availability of resources

These different school climate classifications of domains serve to illustrate how multidimensional and multifaceted the concept of school climate is. However, at the same time it should be noted how difficult is to build such a classification in an exhaustive way. Moreover, these different classifications overlap with each other, and include overlapping constructs within them. Classifications systems risk displaying family resemblance (in the Wittgenstein, 2009, sense). That is, different school climate factors and domain variations, are not necessarily classified within a domain by an exclusive and unique feature. These classifications resemble porous sets, which group different constructs by overlapping similarities. Thus, is not rare to find cases where constructs sit on the edge between two different domains.

Measurement of school climate in large scale assessments

Given the measurement sophistication in large scale assessments within major international surveys, a natural expectation would be to use these studies as a source for understanding relationships with school climate. Nevertheless, the many different large scale assessments echo the state of diversity of school climate research we have previously reviewed. All of these studies assert in their reports the common phrase of 'school climate has been recognised as one of the most influential factors on educational outcomes' – this appears to be a common assumption in large scale assessments sponsored by the OECD: Program for International Student Assessment (PISA) (OECD, 2010a), Trends in International Mathematics and Science Study (TIMSS) (Mullis, Martin, Robitaille, & Foy, 2009) , PIRLS (Mullis, Martin, Kennedy, & Foy, 2007), Teaching and Learning International Survey (TALIS) (OECD, 2010b) and International Civic and Citizenship Education Study (ICCS) (Schulz, Ainley, Fraillon, Kerr, & Losito, 2010). However, all of these studies use different measures, reflecting different emphases on what counts as school climate, whether this involve student-teacher interaction (TIMSS), teacher commitment (PISA), sense of belonging and open classroom discussion (ICCS), or quality of social relations (TALIS). All have in common an emphasis on the study of classroom setting and school context over socioeconomic factors, by means of students' perceptions, teachers' perceptions or principals' reports. As other authors have noticed, in spite of the varying ways in which school climate have been outlined in research, studies have typically reported positive relationships with students' achievement (Collins & Parson, 2010) as well as other outcomes.

School climate and other overarching concepts: school culture and school ethos

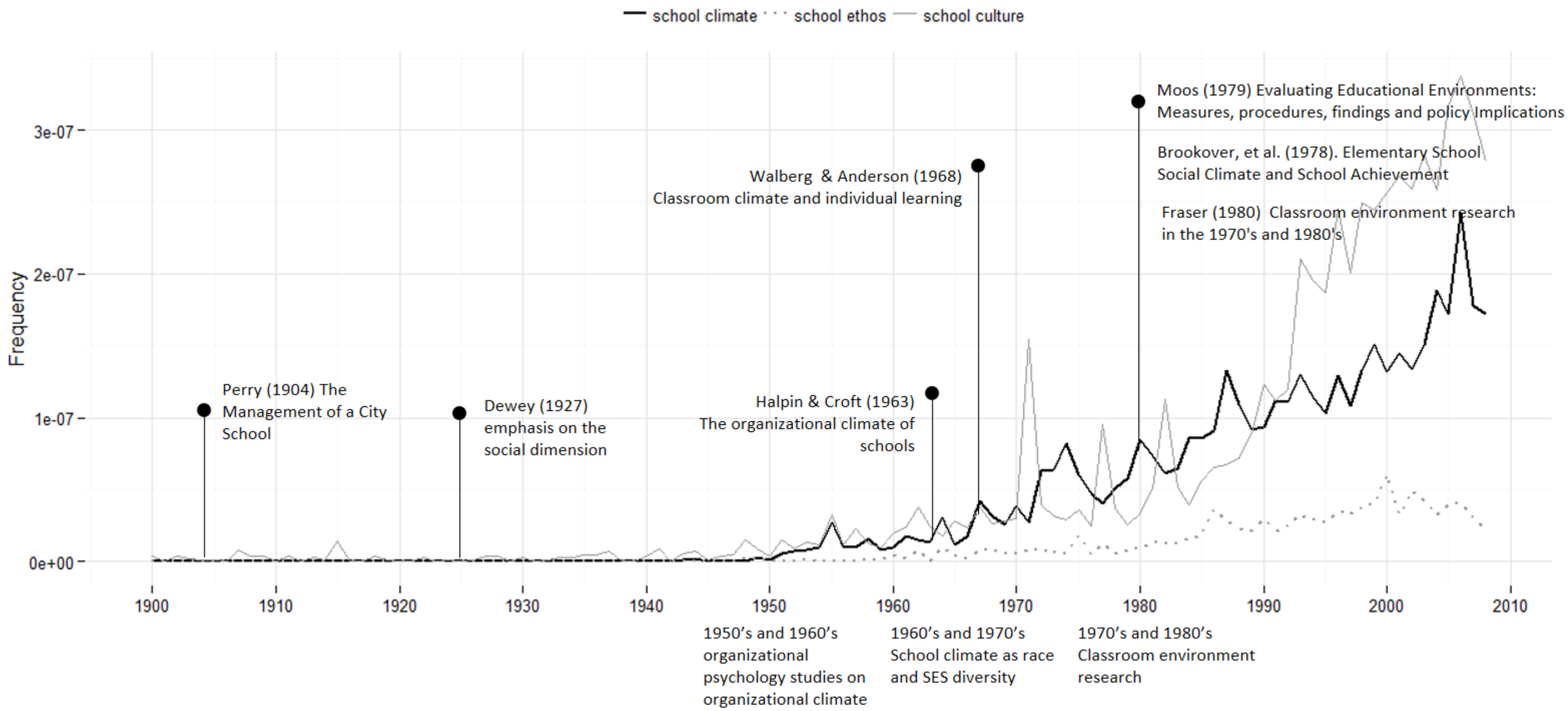
It should be noted that although school climate, culture, and ethos are often used inter-changeably (Glover & Coleman, 2005; Solvason, 2005), some nuanced conceptual distinctions can be made. School culture typically emphasises ritual, beliefs, practices, values and norms (Maslowski, 2006), and sometimes the use of typologies of organizational cultures (Handy & Aitken, 1986). School ethos is more commonly used in British reports and articles, but Solvason (2005) suggests that school culture brings school ethos as a product: while a school may *promote* a given culture, students may *feel* a certain ethos or ambience. Thus, it is common to find articles with these notions together in the titles (Roach & Kratochwill, 2004; Ashkanasy, Wilderom, & Peterson, 2001; Cohen, 2012; Lindahl, 2006).

Van Houtte (2005) has suggested that school culture is much better understood as a component of school climate, a distinction that might be considered congruent with the multidimensional classification of school climate domains (Cohen et al., 2009; Thapa et al., 2013; Wang & Degol, 2015; Zullig et al., 2010). School culture may be seen much more an anthropological concept in nature, and more suitable for the descriptions of cognitive aspects of groups (values and norms). In contrast, school climate is more congruent with a psychological tradition, especially from its organizational roots. Therefore, it may be seen as more sensible to keep this concept as a depiction of the whole school (the whole organization). In contrast, Schoen and Teddlie (2008) propose a contrary approach, in which school climate is a level of school culture. Based on the concept of organizational culture from Schein (2004), the authors propose to keep this hierarchical distinction, by which school climate are 'espoused beliefs', which are a product of the more fundamental basic assumptions (school culture).

A ngram graph retrieved from Google digitalized books database[1], that contains a sample of 4% of all English written books to date (Michel et al., 2011), shows the trends in which these three notions have been used in different books over time (see Figure i1.3). In general, in English-language books, school climate is a more frequent term than school ethos. School culture and school climate were closely related until the 1990s, a time point where 'school culture' became increasingly used in publications.

[1] This plot was produced using the ngramr R package (Carmody, 2014). A web app is also available here: http://books.google.com/ngrams This utility peaks a unit of frequency for every book which contains 40 times the requested sequence of words, an ngram. For example, "school culture" is a sequence of two words, school and culture.

Figure i1.3 Google ngram trends of ‘school climate’, ‘school ethos’ and ‘school culture’ in English-language books published between 1900 and 2000



Given the several definitions that have been attached to these terms, we should regard all three of them as umbrella concepts (Jackman, 1985). Their overlapping meanings are the product of the variety of uses in different studies and different research traditions. In the present research, we therefore start our empirical work from the position that school climate is not a single tightly-defined construct, but rather ***an umbrella concept that allows us to group together numerous factors related to the school experience of the multiple stakeholders (staff, pupils, parents) in a school community.***

It is worth noting that the challenge of pinning school climate down in terms of empirically measureable constructs is greatest when it comes to the social aspects that have been emphasised in so much of school climate research. Whereas factors such as time use or objective features of the school/classroom environment seem easy to understand and can be defined more or less un-controversially, many of the key school climate factors implicated in research are social factors that are much more subjective in nature and thus more difficult to access (C. S. Anderson, 1982; Cohen et al., 2009; Fraser, 1989; Zullig et al., 2010).

However, it would be inappropriate to reject notions such as ‘cohesiveness’ as non-existent, just because they are difficult to conceptualise, define, and/or measure. For example, the work of Bryk and colleagues (Bryk & Schneider, 2004; Bryk et al., 2010) shows how the different levels of ‘trust’ in teachers are related to school achievement gains under the school reforms; similarly Hoy (2012) identified the same factor as positively related to school achievement over and above students’ socioeconomic background. These are refined school climate factors which resonate with the early notions of emotional tone and atmosphere present in the work of Perry (1908), Withall (1949) and Brookover (1979).

Clearly, these subjective, social dimensions of school climate are playing some kind of important explanatory role in influencing student achievement and other school outcomes. But in order to understand this better, these dimensions need to be understood in terms of theoretical frameworks.

Chapter 2

Issues in school climate research and model specification

School climate research presents a range of challenging issues that researchers need to address and resolve (Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015). However, the best way to deal with these issues is not always clear cut. In this chapter, we describe the most common issues within the school climate literature identified in previews reviews (C. S. Anderson, 1982; Cohen et al., 2009; Freiberg, 1998, 1999; Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015; Zullig et al., 2010), we identify the fundamental problem underlying these issues, and we propose ways in which these issues can be resolved.

I summarize these issues under eight headings: a) multidimensionality; b) nesting; c) inference level; d) referent; e) complex relations; f) common method bias; g) causality; and h) temporality. All in all, these issues can be considered research gaps in the literature, because even though these are topics acknowledged by the previous literature, current recommendations are deemed partial, incomplete or absent. Nevertheless, we acknowledge, the current chapter could not be possible without the large contribution of previous research and reviews on school climate research (C. S. Anderson, 1982; Cohen et al., 2009; Freiberg, 1998, 1999; Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015; Zullig et al., 2010). The current chapter aims to move this research programme forward by describing first what these issues are and reviewing alternatives strategies for addressing these problems.

Issue of multidimensionality

The non-exhaustive nature of school climate (Freiberg, 1998) can easily lead to confusion for any systematic review on the topic, given the variety of labels that are attached to similar constructs, and the overlapping nature of different school climate dimensions. Cohen and colleagues (2009), Zullig and colleagues (2010) and Wang and Degol (2015) do converge in the existence of four similar domains: safety, academic, connectedness, and environment/facilities. Yet even here, matters are not straightforward.

Let us take for example the domain of connectedness from Wang and Degol (2015). Within this domain, constructs of school belonging, school connectedness, and school bonding, are all different forms of emotional engagement with the school (Wormington, Anderson, Schneider, Tomlinson, & Brown, 2016). Nevertheless, all these constructs, albeit similar and overlapping, come from different research roots. This comprises an additional effort for researchers, to either know this commonality in advance, or to find out during a literature review process, that these similar construct with different names, may contain relevant findings for their study. All of this carries the risk of missing one or a few different labels, and as consequence to ignore relevant work in the field of interest.

In a similar way, school discipline can be categorized within either academic or safety domains depending on the outcome of study. Sandoval-Hernandez, Aghakasiri, Wild, and Rutkowski (2013) used indicators of school discipline (i.e., classroom disruption) as an indirect way to isolate effective teaching time and predict academic achievement. This is different from the use of school discipline as a safety measure, which was the intended use of the indicator of school discipline policy in Rutkowski, Rutkowski, and Wild (2013). In this latter study, the authors aimed to investigate school factors related to bullying, and the school discipline policy was used as a covariate in this regard. Similar cases can be made for other school climate factors and different domains.

The fundamental problem that lies behind the multidimensionality issue in school climate research is the translation of the conceptual nature of school climate into an all-encompassing construct conception. It is a categorical mistake, in the Rylean (Ryle, 1949) sense, to equate a concept for a construct (Markus, 2008). Constructs are theory dependent and correspond to dimensions of empirical variation within a defined population (Cronbach & Meehl, 1955); whereas concepts do not have such a burden. As we argued earlier, school climate is an umbrella concept, in the Jackman’s sense (Jackman, 1985). School climate domains from the different reviews we described in the previous chapter can categorize different constructs in a conceptual fashion. By necessity, these classifications are done via family resemblance (Wittgenstein, 2009). These domains classify different constructs and aspects of the school experience in terms of overlapping similarities.

This view helps to explain why prescriptions regarding school climate seem ambiguous or vague. For example, ‘positive <u>school climate</u> is an important component of effective schools' (Koth, Bradshaw, & Leaf, 2008) is not a clearly interpretable assertion for school climate research. When this expression is said, we do not know which of the multitude of possible factors from Cohen and colleagues (2009) or Zullig and colleagues (2010) is the critical factor in use which makes that sentence an empirical assertion, an affirmation of the actual state of affairs. But this is not the case when we interpret school climate as an umbrella concept which groups together different constructs. With this assumption in mind, we can then begin to translate the assertion into something similar to 'positive characteristics of the school environment are an important component of effective schools', or more concretely, 'higher levels of classroom discipline and lower levels of classroom disruption are positively related to schools' academic achievement".

I believe that recognising that school climate is not a uni-dimensional construct but a multi-faceted umbrella concept helps rather than hinders us in moving towards a theoretically-rooted account of different school aspects and processes. The higher the abstract level of a school climate dimension the less clear its relation to an outcome. Thus, all-encompassing abstraction may lead to vaguer model specifications. Model specifications are abstract ways to formalize our conceptual models into statistical estimations, in order to answer our research questions. The vaguer the conceptual model we use, the less clear the best model specification is for a given research question.

To summarize, in general terms, school climate is not exhaustively defined by a single construct nor by a list of these. Instead, it is a concept that englobes several constructs. These are often high inference measures (Chávez, 1984), frequently collected with surveys (Wang & Degol, 2015) to capture shared perceptions among different members of the school, about a school or classroom referent (Marsh et al., 2012), that aim to describe the school or the classroom scenario/context.

Issue of nested data or clustered observations

A common issue within school outcomes research, is the nested structure of the observations. This is pervasive for measures of school climate factors (Marsh et al., 2012; Wang & Degol, 2015), especially for measures collected from different respondents clustered either within classrooms or schools. The problem with this issue arises from the shared variance among observations from the same cluster. **Ignoring this data structure can lead researchers to commit Type I error, thus advocating**

effects that are not there, and also to misspecify the level of inference, thus making conclusions at the wrong level.

To illustrate this point, let us take the example of a measure of Classroom Discipline, used in the Teaching and Learning International Survey from 2013 (TALIS 2013 for short) (OECD, 2014e). This is a multi-indicator scale of four Likert-type items. Teachers answer the following questions (see Figure i2.1):

Figure i2.1 Classroom Discipline Items (TALIS 2013)

41. How strongly do you agree or disagree with the following statements about this <u>target class</u>?

Please mark one choice in each row.

	Strongly disagree	Disagree	Agree	Strongly agree
a) When the lesson begins, I have to wait quite a long time for students to quieten down.	☐1	☐2	☐3	☐4
b) Students in this class take care to create a pleasant learning atmosphere.	☐1	☐2	☐3	☐4
c) I lose quite a lot of time because of students interrupting the lesson.	☐1	☐2	☐3	☐4
d) There is a lot of disruptive noise in this classroom.	☐1	☐2	☐3	☐4

One can expect to see shared variance among the answers to this question from teachers at the same school, especially for teachers from the same target class. This is because teachers are referring to the same students. Thus, it is expected that observations from the same cluster are correlated with each other, in comparison to observations from different clusters, because these refer to the same referent.

When observations are correlated within clusters, this represents a threat to the validity of linear regression estimates. Linear regression estimates assume independence of observations (Antonakis, Bendahan, Jacquart, & Lalive, 2010; Rabe-Hesketh & Skrondal, 2012). Neglecting this shared variance leads to Type I error inflation (O’Connell & McCoach, 2008; Snijders & Bosker, 2012). This is similar to presuming that we have a higher number of observations to estimate a parameter. At the same time, ignoring the clustered nature of the data, and simply comparing teachers’ individual scores on this variable with an outcome, can lead to drawing conclusions about individual teachers’ experience of classroom discipline and the outcome. This would, in effect, ignore the intended inferential level of comparing between classroom variance of discipline and the outcome under scrutiny.

Let us say we need to estimate the mean of classroom discipline. If we sample at random only one teacher per school, collecting 100 surveys out of a sample 100 schools, we will have 100 independent observations. Let us call this scenario A. Our standard error can be estimated as the ratio of standard deviation of the parameter, over square root of its sample size:

$$Standard\ Error = \frac{\sigma}{\sqrt{n}} \tag{1}$$

Where σ is the standard deviation and *n* is the sample size. In linear regression models, the standard error is a function of the sample size. Thus, augmenting the sample size minimizes the standard error and gives more power to any inferential test. However, if we sample at random 5 teachers per school, and we sample 20 random schools out of a population of 100 schools we have kept the sample size constant. Yet, we have change the sampling mechanism used to retrieve observations. Let us call this scenario B. In this case, the standard error should be estimated, by correcting for the Variance Inflation Factor (VIF) or design effect to account for the data dependency between schools (O'Connell & McCoach, 2008):

$$Standard\ Error = \frac{\sigma}{\sqrt{n/VIF}} \tag{2}$$

The variance inflation factor or design effect (Snijders & Bosker, 2012), is estimated as:

$$VIF = 1 + (m - 1) * ICC \tag{3}$$

Where, *m* is the average cluster size, or amount of collected cases per cluster, and ICC is the Intra Cluster Correlation coefficient. This is a measure of data dependency. It informs us how much variation in a given variable is shared *between* clusters, instead of *within* clusters. It is often interpreted as the expected correlation out of two observations within the same cluster. For nested data, it follows, that the amount of cases necessary to augment power for inferential test is not a constant; is also a function of how much dependency between cluster the data has (Snijders & Bosker, 2012).

In practical terms, a simple regression analysis and other linear general models, would assume we have 100 independent observations in both scenarios A and B, being blind to the different sample mechanism present in scenario B. Thus, it assumes the same power size for both scenarios. In contrast, a multilevel model approach can account explicitly for the data generation process (Rabe-Hesketh & Skrondal, 2012), and correct standard errors for each scenario A and B, by partitioning the variance between school observations and teachers.

Because of the inflation of Type I error of linear models over nested observations, most of the reviews recommend the use of multilevel models to account for the complex data structure on school climate factors and appropriate level of inference (Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015). However, what is often ignored in the literature is the existence of other methods that also account for data clustering, such as clustered errors, clustered sandwich estimators, Taylor Linearization among other variance estimations methods (McNeish & Stapleton, 2015). These other estimation methods, besides multilevel models, are a complementary addition for special scenarios to work within school climate research. Especially useful when there are problems of low sample size at cluster units, or high data sparsity (large amount of singletons, or clusters with only one case), and also of interest when the outcome is a binary variable (McNeish & Harring, 2015; McNeish & Stapleton, 2014). These methods, are limited in contrast to multilevel models, because these do no provide the same amount of parameters and rich information. However, these are still useful to estimate contextual effects (Asparouhov, 2006; McNeish & Stapleton, 2015; Stapleton, 2013). For example, Caro and Mirazchiyski (2012) used this approach to estimate compositional effects of socioeconomic status on reading and literacy in several countries.

The main difference between design based methods (e.g., clustered errors, sandwich estimator, Taylor Linearization) and model based methods (e.g., HLM, MLM, mixed methods) relates to the meaning of parameters (McNeish & Stapleton, 2015). Because multilevel models partition the variance into within and between components, the estimates are said to be cluster specific. In practical terms, the relationship between the covariate and outcome of interests is estimated within each cluster, and variance terms are obtained to quantify the distance from the grand mean, and from the cluster means. This enables inference test for both type of effects: between and within parts. In contrast, design methods aims to correct the standard errors, and provide population estimates (Snijders & Bosker, 2012). These methods correctly estimate the relationship between covariates, and quantify these relations in terms of distance to the grand mean only, without partitioning the variance. Thus, the estimated effects of these models, can be interpreted as difference between observations as a whole, without specific inferences between clusters. The advantages of each method must be judged with attention to research question and research design at hand.

In summary, given the nested structure of the data of school climate measures, it is essential to correct estimates for data dependency within clusters. At least two approaches are available for this endeavour. The most common approach is the use of model based models (multilevel models, mixed models) and other variants such as multilevel SEM (Marsh et al., 2012). For special case scenarios, the use of design based methods (clustered errors, sandwich estimators, Taylor Linearization, Bootstrapping, etc) (Asparouhov, 2006; McNeish & Stapleton, 2015; Stapleton, 2013) can be recommended.

Issue of inference level

It is important to note that none of the definitions reviewed in chapter 1 have catalogued school climate factors as a within-individual characteristic; the emphasis is unfailingly on higher-level analysis of classroom or school differences. However, it is not uncommon to find articles where classroom level properties are specified as single level factors, without any aggregation. Marsh et al (2012) gives such an example from Ciani and colleagues (Ciani, Middleton, Summers, & Sheldon, 2010, p. 97):

> […] when attempting to predict between-class variance in a motivational outcome variable with classroom goal structure, researchers often use the classroom aggregate of student perceptions of classroom goal structure as the level 2 predictor variable. While not an entirely inappropriate technique, we believe that student perceptions are better suited as level 1 predictors, whereas teacher ratings of classroom goal structure and instructional practices are best suited as level 2 predictors (e.g., Anderman & Young, 1994; Urdan et al., 1998).

This is entirely different from the practice of aggregating responses of students regarding a classroom level construct. Godfrey & Grayman (2014, p. 1806) gives a good example:

> Student-level scores were aggregated across students within each classroom to create an average classroom-level score for each classroom where higher scores indicate higher perceptions of open classroom climate. Aggregations of

> individual student perceptions were used to represent the openness of the classroom climate at the ecological level at which it is conceptualized and to reduce the threat of same-reporter bias (Chan 1998; Godfrey and Yoshikawa 2012).

There is contentious discussion regarding the proper level of analysis for school climate research (van Horn, 2003; Koth et al., 2008; Burkhouse, 2009), as individuals' perceptions of school climate could be important for shaping their respective educational outcomes, as much as or even more than the school-level aggregates of school climate perceptions. For example, Koth and colleagues (Koth et al., 2008) report that individual differences explain a large portion of the variance in responses to items about school climate (65% and 86% of explained variance). Such variations are unsurprising in view of the wide range of individual attributes that could influence perceptions of the school experience. The key point is that when the perceptions are aggregated at the school level, the resulting school-level variations do appear to help us account systematically for differences in school results and other outcomes. This is especially the case for cluster level means parameters, either at the school or classroom level. Other authors have a much more assertive view, and dismiss the use of residual ratings for cluster level constructs (classroom or school referent) (Marsh et al., 2012).

Griffith (2000) showed that the perceptions of school climate aggregated at the school level are related to school outcomes (satisfaction with the school and academic performance). Multilevel analysis over the same data, while controlling for characteristics of the respondents, show similar results (12 out of 16 effects present the same direction). This as an example of how cluster level aggregation of measures does predict cluster level outcomes. Van Horn (2003) showed that a slightly modified version of the 'School Climate Survey' (Kelley et al., 1986) has a stable factor structure within and between schools, and that its scales do predict school level variation in children's academic outcomes. Marsh and colleagues advocate the use of doubly latent models (Marsh et al., 2009, 2012; Morin, Marsh, Nagengast, & Scalas, 2014). In these measurement models, indicators are responded to by school members and describe school or classroom level properties. Then, the responses to these indicators are modelled within multilevel factor analysis, fitting invariant measurement properties for all respondents (within), and for all clusters (between).This approach has advantages over traditional multilevel approaches (Lüdtke et al., 2008; Lüdtke, Marsh, Robitzsch, & Trautwein, 2011), given that this alternative can jointly control for measurement error and sampling error (Televantou et al., 2015), both of which are threats to the estimation of contextual effects (Lüdtke et al., 2008; Pokropek, 2015).

The choice of aggregation to classroom or school level can vary across studies. As noted earlier, Griffith (1995) showed that the appropriate level of inference for school climate depends on the research question and research design. In some scenarios, both levels can be validly addressed, if data is available for students, classroom and schools. However, **the fundamental problem that lies behind the inferential level issue is the problem of inferential level reduction.** This problem can manifest in two ways: as the **ecological fallacy** or as the **atomistic fallacy** (Rabe-Hesketh & Skrondal, 2012). When observations present a hierarchical structure, thus observations are nested in clusters, which is especially the case for students in classroom and schools, and also for teachers in schools, aggregating measures and conducting regressions over school level indicators can lead to the idea that these relations are equal to what would have happened between observations. This is known as the ecological fallacy: aggregated indicators are assumed to represent within relations, and thus between relations are equated to the former. In practical terms, this equates the relation between schools as a good representation of what happens between students.

The atomistic fallacy is its counterpart. In this scenario, relations between observations are assumed to be replicated at the aggregated level (i.e. the cluster), and hence the latter can be reduced to the former. In practical terms, this equates the relation between students as a good representation of what happens between schools. This is why school climate researchers recommend so eagerly the use of multilevel models (e.g., Marsh et al., 2012; Wang & Degol, 2015); these models enable us to separate within cluster relations and between cluster relations.

Provisionally we can say that, in summary, the traditional level of inference for school climate factors is most of the time the cluster level, either the school or the classroom, depending on the construct, how it was measured, and what the research question is. Some exceptions do exist, when both inferences levels could be interpretable. This would be illustrated when formative and reflective measures are discussed.

Issue of referent of indicators

Ideally, most of the school climate factors are cluster level referent measures (Marsh et al., 2012). That is, different measures responded to by several members of the school or classroom groups, via cluster aggregation allow us to characterize a property of the cluster unit (i.e., the school or the classroom). This scenario is also referred in the literature as level 2 reflective measures (Lüdtke et al., 2008, 2011). One of the features of this approach is that at the aggregation level, individuals' responses are interchangeable with each other, as all responses are referred to the cluster level unit of analysis. In the previous example, we used a classroom discipline measure in which the classroom is the referent. If different teachers answered the same questions, referring to the same classroom group, all responses could be interchangeable, assuming no bias: all teachers should coincide regarding the classroom discipline responses, if this is a stable property of the classroom. This is an example of a reflective measure at the classroom level.

However, school climate factors could be measured by level 2 formative constructs (Lüdtke et al., 2008, 2011). Such constructs are built up by aggregations of validly interpretable level 1 indicators. The most common example of this scenario is the school level means of the socioeconomic background of students. This factor is less often included as a school climate factor in its own right, yet is one of the most common examples present in the educational literature of formative constructs. A similar example is present in the study of 'big fish little pond' effects (BFLP). In these studies, the students' achievement is aggregated at the school level, to capture the between school variance in achievement, and estimates its relations to academic self-concept. This is a paradoxical effect, in which students in schools with higher achievement have lower levels of self-academic concept, in comparison to students with similar levels of ability, in schools with lower achievement levels (e.g., Marsh, Kuyper, Morin, Parker, & Seaton, 2014; Marsh, Trautwein, Ludtke, Baumert, & Koller, 2007; Marsh, Abduljabbar, et al., 2014; Nagengast & Marsh, 2012). Another example is provided by Engel and colleagues (Engel, Rutkowski, & Rutkowski, 2009). These authors, used aggregated levels of responses to a self-report measure of bullying, to produce an indicator of bullying rates at the school level. Then, the authors used the measure to estimate contextual effects of school bullying on math achievement. All these previous cases constitute examples of how level 1 measures are used to create formative level 2 constructs. Marsh, et al (2012) makes a similar distinction, calling climate factors those constructs with level 2 reflective properties, and contextual factors those indicators with level 1 properties which can be aggregated as level 2 formative constructs.

Unlike reflective level 2 measures, formative level 2 measures in multilevel models do produce interpretable within and between estimates. We can illustrate this with an example. Let us say we have the following multilevel model fitted to estimate the relationship between math achievement and bullying. In this scenario, bullying is a binary variable, classifying whether or not a student has suffered from bullying in the last month. The school aggregated average of this covariate expresses what proportion of students have suffered from bullying in the last month. To estimate the relationship between these factors we can borrow the contextual effect model specification from Willms (2010), and fit the following model:

$$Math_{ij} = \gamma_{00} + \gamma_{01} * X_j + \gamma_{10} * (X_{ij} - X_j) + \mu_{0j} + e_{ij} \tag{4}$$

The first parameter, γ_{00} is the school expected mean when schools and students present average levels of bullying, when X_j is centred at the grand mean and X_{ij} is centered at the cluster mean (Ma, X., Ma & Bradley, 2008). This is different from the classical specification from Blalock (1984), in which not centering is specified. Given the use of this centering within cluster, we explicitly state $(X_{ij} - X_j)$ in equation 4 for parameter γ_{10}. The parameter γ_{10} estimates the difference of being bullied or not, between students clustered in the same schools. These are pooled estimates across all schools (fixed slopes). Moreover, parameter γ_{01} is the expected mean difference of math achievement between schools, for every unit change of X_j. If this covariate is scaled in percentages from 0 to 1, it is the expected difference in school math achievement for schools, between 0% and 100% of bullying rates. If we subtract parameters $\gamma_{10} - \gamma_{01}$, we get the contextual effect. If these two parameters were equal $\gamma_{10} = \gamma_{01}$, this would suggest we gain no information by adding the cluster level means of the covariate (Algina & Swaminathan, 2011). If the contextual effect is different from zero, $\gamma_{10} - \gamma_{01}$, this represents the expected mean difference between schools, due to specific variations of bullying rates, regardless of students' individual experiences of bullying. This model is statistically equivalent to the grand mean centering specification, for fixed effect model, i.e. random intercepts and fixed slopes (Algina & Swaminathan, 2011; Enders & Tofighi, 2007).

The interpretability of within and between estimate coefficients is not controversial with formative level 2 constructs. This is not the case with reflective measures. Marsh and colleagues (Marsh et al., 2012) considered it to be problematic to interpret within estimates in these cases, following Cronbach's (1976) assertion that the perceptions of individuals, i.e. informants of cluster level constructs, are an entirely separate problem from the measurement of environments, i.e. cluster level features. Thus, the interpretation of within estimates in multilevel models, the parameter γ_{01} in the previous example, can be regarded as non-traditional, for reflective level 2 constructs, in comparison to formative level 2 constructs. Nevertheless, whether to interpret this parameter or not should be left to theoretical judgment.

Moreover, indicators can sometimes be formulated as reflective or formative by changing the referent. Wang and Degol (2015) illustrate this with following example: "I feel I fit in at this school" (formative form) and "Students feel that they fit in at this school" (reflective form). Similar conversion between items can be made for bullying indicators, thus permitting to create either reflective or formative level 2 constructs. Nevertheless, these can lead to different findings.

One of the risks within school climate research, for formative and reflective measures of school climate factors, is an issue of endogeneity. If the within part of the covariates is not included in the model, then the between effect of the aggregated level 2 variables would be biased, and might overestimate the effect of the school climate factor. In order to correct estimates for this issue, the recommended strategy is to include the covariate and its cluster means in the model (see Antonakis et al., 2010; Antonakis, Bendahan, Jacquart, & Lalive, 2014 for a revision). Campbell (2008) applied this method of including level 1 and level 2 forms of the index of classroom discussion. This yields corrected estimates of the level 2 construct, and allows estimates of contextual effects (i.e. the between effects independent of individual variations). This is different from the practice in Godfrey and Greyman (2014), which only included the aggregated form of the school climate factor.

Additionally, as Campbell (2008) notes, the inclusion of level 1 and level 2 forms of the index will lead to noise within the model, if appropriate centering techniques are not used. The researcher must purge any correlation between the individual level indicator of school climate factor from its aggregated form. This is achieved by centering the school climate factor to the cluster mean, and centering the cluster means to the grand mean (Algina & Swaminathan, 2011; Enders & Tofighi, 2007). This specification divides the covariate into two perfectly orthogonal components: the within and between parts. The first expresses the difference of the observations from the cluster mean, whereas the between part expresses the relative differences of all clusters to the grand mean of the covariate.

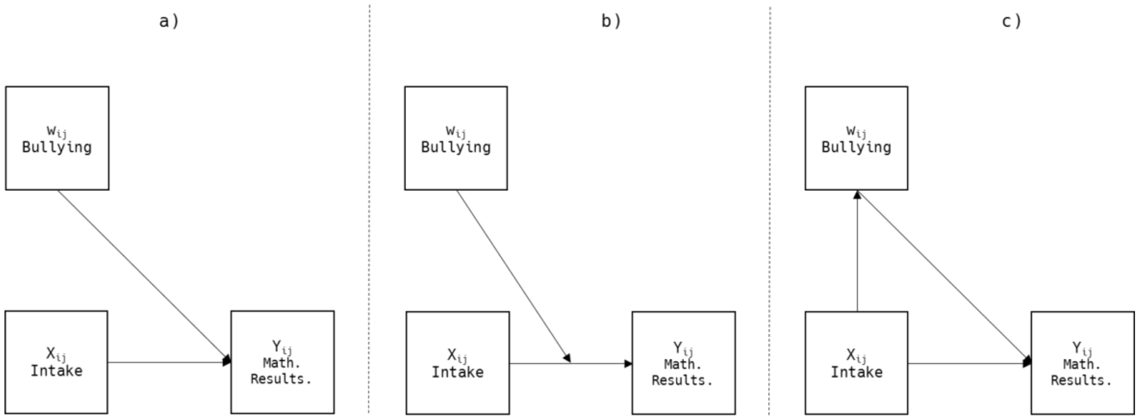
Much of the emphasis in classroom and contextual effects estimation has relied on the use of multilevel approaches (Marsh et al., 2012; Wang & Degol, 2015), and respect for the appropriate level of inference. However, there is less guidance in the literature regarding the rationale and appropriate specification for covariates as within and between components in mixed models. The partition of covariates into its within and between parts, and the use of an appropriate centering technique are crucial.

Issue of complex relations

School climate factors are in most cases descriptions of features of the learning environment. As such, these are fixed effects, to both outcome and covariate variables. Thus, these features can be systematically linked to the outcome variable, the covariate variable, and/or the relationship between covariates.

We can try to illustrate this with three different scenarios. These are depicted in Figure i2.1.

Figure i2.2 Different relations between a school climate factor, a covariate and an outcome.



Let us picture plausible relations between three variables: student intake characteristics (a covariate), students' math achievement (an outcome), and bullying rates at the school level (a school climate factor). Just for the sake of the example, let us ignore the difference of within and between components of the school climate factor, and keep our attention on the possible relations among all of the variables.

In scenario a) we specify bullying as having a direct effect on math results, and no correlation with student intake characteristics. A traditional multilevel model has this assumption, and this model is not problematic for traditional model estimations. Scenario b) specifies that the relation of student intake characteristics and students' math results, are conditional to the level of the school climate factor, in this case to the level of bullying at school. This model can be specified within a multilevel framework as a cross level interaction, where the slope of regressing the outcome on the covariate varies according to different values of the school climate factor (O'Connell & McCoach, 2008; Rabe-Hesketh & Skrondal, 2012; Snijders & Bosker, 2012). However, scenario c) is an indirect effect scenario. It postulates that at least a portion of the relationship between student intake characteristics and the outcome occurs thanks to its relation to the school climate mediator, bullying. This scenario is problematic for traditional multilevel models because there is no way to estimate relations between covariates of a model. These are always assumed to be independent. Thus, mediation models or indirect effects can hardly be estimated with traditional multilevel analysis, unless odd assumptions and complicated workarounds are used; even there, no inferential tests for the indirect effects are produced (see Liu, Van Damme, Gielen, & Van Den Noortgate, 2015 for an example). It is also plausible to conceive mixture scenarios, in which b) and c) jointly exist. These three scenarios are just examples, as more complex relations between these variables could be possible.

The **fundamental problem of the issue of complex relations is the overreliance on traditional methods of estimation of effects and its limitations for research as a consequence**. "[…] it is tempting, if the only tool you have is a hammer, to treat every problem as if it were a nail" (Maslow, 2002, p. 15). The limitations of traditional methods and approaches may lead to relevant questions being left unanswered, because of the limitations of the statistical models at use, not because of the limits of plausible conceptual models. Multilevel SEM models are a desirable alternative in this regard, because this method accounts for the nested structure of the data, produces valid inferences at each level, deals with indirect effects estimations (K.

J. Preacher, Zhang, & Zyphur, 2011a; K. J. Preacher, Zyphur, & Zhang, 2010) and also allows to test interaction effects (Ryu, 2015; Zyphur, Zammuto, & Zhang, 2016).

The three remaining issues are not unique to school climate research. Concerns about common method variance, causality, and temporality are more general problems, shared with other applied disciplines besides educational research that rely on the use of self-administered surveys and tests within a cross-sectional design, such as sociology, social psychology, cross- cultural psychology, and organizational research, among others.

Issue of method variance

The issue of common method variance is raised where the same instrument type is used to collect observations. However, in school climate research, this translates more concretely to the fact that students describe their experience within their school and give responses on psychological variables and/or behaviour, and – at least in some cases – the latter are also the source of the outcome under investigation (Wang & Degol, 2015). In the organizational literature, the attention placed on common method bias is larger than in school climate research. The paradigmatic example of common method bias is where subordinates rate leaders (i.e., supervisors) on two different features that are hypothesised to be linked to each other: for example, leadership and effectiveness (Antonakis et al., 2010). Given the contractual relationship between subordinate and leader in work environments, the scepticism over same source informant, is larger than for the case of students and classroom environments. However, the relationships between school principals and teachers share similar features to those commonly described in organizational literature between leaders and subordinates.

Common method variance is thought to bias estimates of associations between two variables as this source of variance may exert influence on both variables together. Wang and Degol (2015) believe this biases estimates towards the detection of significant effects, by estimate inflation. In contrast, other authors have shown (Antonakis et al., 2010, 2014; Luchman, 2014) that the direction of the bias can go both ways, upwards and downwards deviation of estimates. There are no statistical tricks to purge estimates of this bias. Statistical remediation of bias estimates requires the researcher and/or reviewers to know in advance, plausible omitted covariates with known effects (Luchman, 2014). This condition can be satisfied more easily for variables with a richer history of published studies. Thus, it plays as a disadvantage for newer research. Thus, the control of common method bias entails its consideration in the research design, either by manipulating the methods, using alternative informants, including more informant sources (Wang & Degol, 2015), or by using instrumental variables, purging estimates of common method bias, and assuring consistent estimates (Antonakis et al., 2010).

For the case of school climate research, this is an area that needs more research. It seems plausible that this is more of threat for school effects studies in which the relationship between school principal and teacher is at the core of the study. This last scenario fits much more what is known in the organizational literature on the topic. For example, teachers may feel "motivated" to highly rate their school satisfaction and their views of the school principal, given their contractual relationship. In contrast, research investigations into associations between school climate and objective measures of outcomes (e.g., student achievement results or teacher turnover rates) are less likely to be affected, because the climate factor does not include the same informant under the same method of response. As a consequence, there is no expectation for these two factors to shared common method variance. Nonetheless, given the consideration of complex model specifications (mediation and moderation effects) discussed earlier, we must acknowledge that common method variance could still play a role in biasing estimates of associations among the key school climate factors of interest and other covariates measured by reports from the same informants.

Issue of causality

The problem of causality is a general problem to all disciplines that relies on the use of non-experimental data. Given that non-experimental data does not have observations in which covariates are totally exogenous to the outcome under study, it is difficult to stablish whether a given covariate is the sole cause of an effect. In contrast, in the case of experimental designs, given that the "treatment" is manipulated and is exogenous to the assignment of experimental groups, to estimate the causal effect is straightforward, and a simple linear regression is sufficient. This is not the case for cross-sectional and correlational studies, for which the common mantra says "correlation is not causation" (Antonakis & Lalive, 2011), and thus assumed causal interpretation of estimates is not recommended, and often actively forbidden.

In this regard, most of the reviews in school climate research advocate for the use of longitudinal studies to shed light on the direction of effects between covariates (Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015), under the common assumption that temporal precedence between cause and effect can be illustrated with cross lagged designs. For example, Guay, Marsh, & Boivin, (2003) used a three wave design to assess whether academic achievement influences self-concept, or if the direction of effects goes from self-concept to achievement. These two covariates are the core of the big fish little pond contextual effect (e.g., Marsh, Kuyper, et al., 2014). However, cross lagged designs for cluster level effects of school climate characteristics seem to be an open avenue for research.

A less common recommendation in the literature is the use of causal inference methods (Antonakis et al., 2010, 2014; Murnane & Willett, 2011). This is an absent recommendation in the reviews of school climate research (Cohen et al., 2009; Marsh et al., 2012; Thapa et al., 2013; Wang & Degol, 2015). However, if certain assumptions are met (see Imai, King, & Stuart, 2008), causal inference under the potential outcome framework (Pearl, 2012) can be applied to estimate certain effects. For example Ponzo (2013) used propensity score matching to estimate the casual effect on achievement of being bullied. With this method, Ponzo (2013) estimated the difference in achievement between two students of similar characteristics, thus providing the most informed guess of what its achievement would be like, if a student who was not bullied receives the treatment of being bullied. This approach is not so different from asking what would be the potential "achievement" of student *i* if attends school *j*, given that school *j* has certain school rates of bullying events. Thus, the same question could be plausible for other school climate factors.

The most important message for school climate research is that there are frameworks for the study of causality between variables other than experiments. This is especially relevant when experimental designs are less plausible within school settings. However, this is not a justification for the abandonment of randomized trials designs, when these are justified and feasible. This is a call for the thorough use of secondary data of educational phenomena that is already available. The use of causal inference methods is an open area for further development within school climate research.

Issue of temporality

One of the main assumptions of school effectiveness research is the idea that "school climate" is a malleable property, and thus subject to change. Nevertheless, studies that track the longitudinal trajectories of cluster level properties are lacking. This is different from studies in which the outcome's temporal trajectory is under scrutiny. Research of the latter form do exist. For example, Bryk and Schneider (2004) estimated the relation of teachers trust to school learning gains, and reported positive results; Petras, Masyn, and Buckely (2011) estimated which students were at higher risk of school removal and estimated the effects of classroom level aggression, in addition to students' aggressive behaviour. These authors found that students in more aggressive classroom are less likely to be removed from schools, in spite of displaying a 'risky profile'. Research designs closer to the idea of capturing school climate change in time were presented by Bryk and Schneider (2004). By using a multivariate multilevel model, the author estimated the relation between teacher trust changes and teacher innovation.

Research into how school climate factors may change over time are relevant, as these effects may influence outcome trajectories and their relations to developmental trajectories (Wang & Degol, 2015), such as learning growth, school adjustment during school transitions, social attitude formation, youth health outcomes (e.g., sexual activity, substance use, risky behaviours), among other outcomes of interest. All of these may also interact with whole school interventions in any of these areas, as program adoptions and fidelity may well be captured by school climate type factors (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) or influence outcomes indirectly via school climate factors (Banerjee, Weare, & Farr, 2014).

In summary, longitudinal patterns of school climate factors or clustered level effects are research avenues open for exploration. The outcomes of such research are relevant for public policy, whole school program interventions, and the study of developmental trajectories. In general, these are applicable to any outcome trajectories study. If school climate factors prove to be relevantly different between schools for particular outcomes, it suggests these relations are potentially relevant for changes across time for schools as well.

The present research

The overall research aim of the current thesis is to address the question, “What is the relationship between school climate and school outcomes?” However, as we have argued earlier, school climate does not represent a singular dimension, nor is this the case for school outcomes. Therefore, to address such a question, we need to frame the overarching aim into more concrete questions. This entails proposing research questions with specific school climate factors and specific school outcomes. Moreover, this requires a design and method that can provide answers for such a questions. In the next section, we summarize the main tenets of the conceptual framework that have guided our studies, along with its general methods.

Conceptual Framework

Given that the present work focuses on the general problems of school climate research, and how to address these, it is better to divide the conceptual framework that guides this research in two parts: its meta-theoretical assumptions and the theoretical models. The first part refers to general assumptions regarding the nature of school climate factors, regardless of the theories about the outcomes. These assumptions are generalized over the four articles. On the other hand, the theoretical models we use are more specific to the outcomes under investigation, and what is known about their predictors and their related contextual factors.

Meta-theoretical assumptions

As noted earlier, school climate is best understood as an umbrella concept rather than a single construct, and it is characterized by a non-exhaustive nature. This assumption allows us to separate relevant clustered level constructs, under the lens of relevant literature of an outcome, to properly specify model estimation. In this way, school climate factors and school outcomes, can be treated as multidimensional and multivariate phenomena. In all the studies we use a school climate factors approach. That is to say, we essentially borrow what is known from the school effects[2] literature,

[2] We use the term effect as in Rabe-Hesketh & Skrondal (2012, p21), without aiming for a causal claim. This choice is used to express the directional nature of the model specifications between two factors, more than its causal interpretation. This is convenient for describing multivariate models, as in this case,

and how to specify models, but we apply it to “high inference measures” that mostly describe subjective experiences from school settings. I use multivariate models to complement traditional model specification, in order to address more complex questions. I avoid the conflation of different constructs onto a single factor of school climate, to represent the whole school experience, as this practice could mask relevant differential effects of each school climate factor separately. Moreover, a fine-grained, differentiated account of specific constructs under the broad heading of school climate allows analysis to test if these factors are inter-related. Hence, complex relations are considered, and indirect effects, moderation effects and moderated mediations are explored where appropriate, on the basis of sound conceptual support.

I assume all variance in school climate scores, derived from responses by individuals within clusters, can be divided into two components, the within and the between parts. The within part of these factors captures deviations of observations from the central tendency of its cluster, whereas the between part captures the differences between clusters. Under centering within cluster, these components are orthogonal and additive (Hox, 2010). Within the latent covariate framework (Lüdtke et al., 2008), the same variance partition is achieved by centering covariates to the latent cluster mean (Ryu, 2015). This partition enables scrutiny of the relationships of school climate factors and other covariates as general relations across a whole sample as well as relations at the cluster level, in a joint manner (Hox, 2010; Lüdtke et al., 2008; K. J. Preacher et al., 2011a, 2010; Kristopher J. Preacher, Zhang, & Zyphur, 2015; Ryu, 2015).

As discussed earlier, I also assume that school climate factors can present themselves as measures of cluster level constructs that take two forms: formative or reflective. As such, these covariates will have different properties which will change the interpretation of the estimates (Lüdtke et al., 2008, 2011; Marsh et al., 2012). Identifying these characteristics is fundamental to the measurement component of the models, and fundamental as well for the conceptual interpretation of the results.

All these assumptions are general and fundamental to the study of how settings and group characteristics are related to individuals and clustered outcomes. In general terms, it is assumed that any conceptual model that expresses the relation between a covariate and an outcome would allow the specification of clustered effects, if either the

---

because we allow for covariates to have directional relations between them, and directional relations to more than a single outcome.

outcome variable or the covariate are known to be clustered. If the covariate of interest is thought to be fixed at the cluster level, and related to all observations in a common way, it can work as a clustered level factor. Here, the distinction between formative and reflective clustered constructs is relevant, as any covariate in a clustered scenario can potentially display a cluster level effect, whether as a clustered construct in principle (i.e. a reflective measures), or as a level 1 covariate in a clustered scenario (i.e. a formative measures). Thus, even if a general theory for an outcome does not encompass cluster level constructs, and does not possess expectations of directional associations between cluster level covariates, it is plausible to assume these relations can exhibit cluster level effects and therefore be worthy of inquiry.

On the non-random assignment to clusters

For the sake of parsimony and abstraction, let us say we are interested in the relation between y_{ij} and x_{ij}, and we know these variables are clustered, hence the subscripts i and j have substantial meaning. These can be outcomes and covariates related either to students or to teachers. In the case of educational outcomes in Chile, this is easy to picture because there are known reasons why students are not randomly assigned to schools, because of system level properties such as school selectivity, school choice and urban segregation (Elacqua, 2012; McEwan et al., 2008; Valenzuela et al., 2014). Similar findings exist for the case of teachers not randomly assigned to schools, because public schools concentrate a higher proportion of less qualified teachers in contrast to private schools and public schools hire more novice teachers in a higher proportion than private schools (Cabezas, Gallego, Santelices, & Zarhi, 2011; Meckes & Bascopé, 2012; Rivero, 2013, 2015; Ruffinelli & Guerrero, 2009). As a consequence, it is plausible to suspect that the relation between y_{ij} and x_{ij}, might be influenced also by clustered level differences as well, which in the case of educational outcomes, pertains to the school environment. School climate factors are a family of different constructs that could describe these potential differences.

Theoretical Models and Research Questions

In the present thesis I primarily use three main conceptual frameworks. In this section, I succinctly outline what conceptual models were used in each study and what research question was addressed in each paper. These models are described in more detail in each paper in the context of model specification for each analysis. Given the broad and multi-faceted nature of the school climate concept, the school outcomes investigated in these analysis are diverse. While a great deal of work on school effectiveness has focused on student achievement outcomes, as discussed earlier, the integration of organizational approaches raises the need to investigate outcomes relevant to key stakeholders. The present work therefore includes studies predicting a broad range of outcomes for teachers as well as students.

The job demands and resources model (Hakanen, Bakker, & Schaufeli, 2006; Prieto, Soria, Martínez, & Schaufeli, 2008) is used in paper 1 and 2, to specify which school covariates and school climate factors are demands and thus positively related to teacher turnover, and which factors are resources, and thus are negatively related to turnover. In general terms, demands are stressor factors that can lead to stress and burnout in work settings, whereas job resources are protective factors, that lead to higher engagement and commitment.

In order to use this framework for the study of teacher turnover (paper 1) and teacher turnover intentions (paper 2), we assume school covariates and school climate factors behave as either risk or protective factors of turnover. Moreover, because teacher actual turnover and turnover intentions are thought to be components of the same turnover and withdrawal process (Podsakoff, LePine, & LePine, 2007), we assume most of the contextual factors may have the same direction of effects on both outcomes, in each study.

In paper 1 we address the general questions of **what is the relationship between school intake characteristics and teacher turnover? and what other factors account for this relationship?** Because school intake composition is thought to be related to teacher turnover in an indirect way, and might involve characteristics of the school environment, we use a school climate factors approach to address this question.

Additionally, in paper 2, we hypothesize that the effects from the protective and stressors factors on the outcomes would be indirect via job satisfaction. The aim of this paper is to answer **what are the relationships between school climate factors, in particular, teacher-student relationships and school discipline, with teacher turnover intentions?** Because teachers are not randomly distributed between schools, and school environments are fixed to groups of teachers, we used a school climate factors effects approach in both papers.

In paper 3, we ask **what is the relationship between bullying and students' achievement?** We integrate models of peer rejection and achievement (Buhs, Ladd, & Herald, 2006; Buhs, Ladd, & Herald-Brown, 2009) with engagement and academic achievement models, drawing upon cluster level variations in students' sense of school belonging (Konishi, Hymel, Zumbo, & Li, 2010; J.-S. Lee, 2014; Norwalk, Hamm, Farmer, & Barnes, 2015; Osterman, 2000). The first portion of the model expects bullied students to reach lower achievement levels due to peer rejection, because of negative effects of the maltreatment resulting in lower participation at school. The second portion of the model, expect negative effects of bullying on school belonging and engagement. As a consequence, we hypothesized indirect effects on achievement of bullying via these belonging and engagement factors. Applying the same principles as earlier, we studied these covariates as school climate factors, because students are not randomly assigned to schools, and it is safer to assume plausible cluster level effects of cluster level properties, such school bullying rates and school belonging.

Finally, in paper 4, we use research on authoritarianism and its relation to prejudice and other social attitudes (Asbrock, Sibley, & Duckitt, 2010; Duckitt & Sibley, 2009) as a basis for the study. Within authoritarianism research, the sophistication hypothesis states that cognitive sophistication is negatively related to authoritarianism (Carvacho et al., 2013). Moreover, Van Hiel, Pandelaere, and Duriez (2004) suggest that educational interventions may be directed to reduce the 'need for closure', a form of cognitive conservatism and closed-mindedness, and thereby reduce authoritarianism. We hypothesize that, higher civic knowledge would be negatively related to authoritarianism, and indirectly related to democratic attitudes. Additionally, we also believed these effects would be conditional upon the level of open classroom discussion between schools. Thus, in this paper, we also question **What is the role of school classroom discussion in the relationship of civic knowledge with ideological beliefs (authoritarianism) and democratic attitudes?** Because open classroom discussion is a school climate factor of a reflective nature, we partition this variable into its within and between components, and scrutinize the role it has for the relationships of civic knowledge, authoritarianism and democratic attitudes.

Methods

Data sources

A Focus on Chile. All the conducted studies include data from Chile, for diverse reasons. The most fundamental reason is that the Chilean educational system constitutes a uniquely informative scenario because it is one of the most unequal educational systems (OECD, 2012), suffering from unequal allocation of qualified teachers (Akiba, LeTendre, & Scribner, 2007) and high stratification of students between schools (Mizala & Torche, 2012). As we have argued earlier, if there is any indication of non-random assignments of observations to clusters, cluster level differences can plausibly be expected. Chile is a good example in which the stratification of both students and teachers is well documented, and I expected more evident effects of school climate factors. Paper 4 is the only paper in which countries other than Chile were included in the analysis, for the sake of testing robustness of the model across varying international contexts.

Two other reasons motivate the focus on Chile, which do not pertain to the school climate research. The research was funded by Becas Chile scholarship, which is designed to promote human capital development. Devoting this research onto Chilean data is a way to maximize the return on the investment placed upon me. Additionally, because I grew up in Chile and previously worked for the educational system, the analysis of school climate factors in the Chilean context was more familiar.

Sample size. The study of school climate factors requires the use of large samples of individuals and clusters, in order to guarantee appropriate power for inferences at the within and between cluster levels (Lucas, 2014). Because my research aims referred to school climate factor estimates, I required large samples of clusters and not only observations from individuals. Estimates of required samples for multilevel models should ideally be no fewer than 50 clusters (Maas & Hox, 2005). Moreover, the requirements to test more complex models, such as indirect effects with clustered data, under multilevel SEM, can be higher than 100 clusters (Meuleman & Billiet, 2009). Given these requirements, I relied on the use of secondary data that fulfil these criteria and address the general problems of school climate research.

Data sources. In these four papers I used data from different sources. In paper 1, study 1, I used data from the ‘Teacher Suitability Survey’ (Valenzuela & Sevilla, 2013), from 2007 to 2013. These are publicly available teachers' contract records from all the schools in Chile. These records were complemented by the Index of Vulnerability of each school (Agencia de Calidad de la Educación, 2012b). In study 2, I used the 2010 cohort from the National Teaching Evaluation System (Santelices & Taut, 2011; Taut, Santelices, & Stecher, 2012), and cross referenced these records with previous sources.

In paper 2, I used data from the Teaching and Learning International Survey from 2013 (TALIS 2013; OECD, 2014), with the subpopulation of Chilean teachers. This is a large scale assessment study, which is also publicly available.

Paper 3 uses the national representative sample of eighth graders from Chile from the Trends in International Mathematics and Science Study 2011 (TIMSS 2011).

Finally, paper 4 uses the data from The International Civic and Citizenship Education Study (ICCS) (Schulz, Fraillon, Ainley, Losito, & Kerr, 2008), including variables from the Latin-American (Schulz, Ainley, Friedman, & Lietz, 2011) and Asian (Fraillon, Schulz, & Ainley, 2012) regional modules. This paper includes data from Chile, Colombia, Guatemala, Mexico, Paraguay, and Dominican Republic, along with data from Taiwan, Hong Kong SAR, Indonesia, Korea Republic, and Thailand.

More details regarding measures and participant characteristics are provided within each corresponding paper.

Issues with secondary data. One of the main advantages of the samples used in these studies is that the results are generalizable to their respective populations, that is, to their sampling frame. This is true especially for papers 2, 3, and 4. Paper 1, contains two studies, and is the only paper which does not use secondary data from large scale assessments. In this paper we use population records in study 1, which provides generalizable results at the census level, whereas in study 2, we use large samples, yet with uncertain generalizability. The second main advantage for the use of secondary data is its zero cost and availability. Primary data collection to arrive at similar datasets would not have been possible due to costs and available resources. However, these advantages do have a trade-off. Unlike first hand experimental studies, in which research questions match perfect research design, here the research questions are constrained by already implemented research designs. Thus, I need to rely heavily on conceptual frameworks that can guide the matching of research questions with conceptual models and analytic model specifications, in order to produce results of interest for the present inquiry.

Analysis Strategies

School climate factors are covariates that are ideally fixed to clusters. However, regarding their type of measurement, either formative or reflective, these can have within and between effects, which need to be accounted for within model specification. This is not the case for school climate factors measured at the cluster level in the first place, which have no variance of any form at the within level. In the next section, different types of effects for school climate factors are described in general terms. Each paper within this manuscript has its own substantial description of the model specification and method of estimation.

Type of clustered effects

Research that dwells on the relationship between individual outcomes and context settings presents different type of effects described in the literature. The most common effect comes from the multilevel approaches that deal with the relationship between context characteristics and individual outcomes (e.g., Feaster, Brincks, Robbins, & Szapocznik, 2011). However, more complex scenarios are also possible, such as multilevel mediations or indirect effects. These models also include, the possibility to accommodate bottom-up effects and top-down effects for the case of individuals clustered in groups (K. J. Preacher et al., 2010).

Feaster and colleagues (2011) argue that the study of how individuals' outcomes are related to group effects can be illustrated at least with three forms of contextual effects: divergence of within and between effects, presence of cross level interactions, and the effects of discrepancies. These type of effects can also be applied to the study of school climate factors. The first type of effects, is commonly known as compositional effects, within the study of educational inequalities (e.g., Caro & Lenkeit, 2012; Willms, 2010). Yet the same model specification is used to get contextual effects of other variables (e.g., Enders & Tofighi, 2007), yielding the same kind of interpretation: a contextual effect is the size of the difference on an outcome, due to the clustered level effect, regardless of the individual clustered deviations. Contextual effects can be obtained with traditional multilevel models, with latent covariate models and multilevel SEM models (Lüdtke et al., 2008; Marsh et al., 2012), and also with clustered errors or design based methods (Asparouhov, 2006; McNeish & Stapleton, 2015; Stapleton, 2013). The critical issue is to partition covariates into within and between components, and get the relevant subtraction of estimates. This is illustrated in papers 2, 3 and 4.

Cross level interactions occur when cluster level covariates condition the relationship between the outcome and another given covariate. Here, an interaction term is used to estimate the conditional relationship between the covariate and a outcome, conditioned by a modifier (Rabe-Hesketh & Skrondal, 2012). However, when variables are partitioned into between and within parts, the interaction could potentially occur between the variable at either part of the model. Ryu (2015) proposes a multilevel SEM framework to address this issue, by suggesting to create the product terms of the interaction of interest between covariates, by either centering within cluster, when the within interaction is of interest, or centering by cluster to the grand mean, when the interaction of interest is at the between part of the model. This is illustrated in paper 3 and 4, although in the latter, we use a hybrid model specification within a design based method approach.

Discrepancy effects are implemented to scrutinise variation in a clustered variable that is related to the outcome. The focus of attention of this type of effect, is to ask whether the size of variability in a factor, hence, lack of homogeneity within clusters of a covariate, is related to an outcome. Thus, instead of estimating the cluster mean, a measure of the cluster variability is inserted in the equation, for example the standard deviation of the covariate for each cluster. Examples of this approach are present in the study of bullying and socioeconomic diversity within schools (Due et al., 2009) and subjective sleep quality and time in bed variability (Wiley, Bei, Trinder, & Manber, 2014). This is one of the type of effects less studied in school climate research. In the work for the present thesis, we did conduct a preliminary exploration of these effects, but no positive findings were found, and thus, analysis of discrepancy effects is not reported in any of the papers in the current manuscript.

Indirect effects within multilevel structures can be estimated under a multilevel SEM framework (K. J. Preacher et al., 2011a, 2010). In this case, covariate, mediator and outcome can all be clustered, and thus the indirect effects between variables can be estimated in almost all possible forms (for illustrative purposes, see K. J. Preacher, Zhang, & Zyphur, 2011b). Top-down effects considered the fact that the covariate or the mediator are cluster level constructs, whereas bottom-up effects considered the scenario where the mediator or outcome are at the cluster level. In a general way, multilevel SEM framework can accommodate each scenario, and to joint scenarios of either form.

These approaches are relevant to the study of school climate factors, because within this framework indirect effects of the factors can be estimated, or because indirect effects occur via school climate factors. This approach is illustrated in papers 2, 3 and 4. Yet again, in paper 4 a hybrid model specification is used to estimate the indirect effects under a design based method approach (Taylor series Linearization).

Overview of the present work

In the previous chapter, we proposed a conceptual origin for the concept of school climate from within school effectiveness research, and reviewed historical approaches to the study of schools contextual factors. In the previous sections of this chapter, we have outlined eight research gaps in the school climate literature, open for further development. In this research work, we address the first five problems in relation to four different settings of school outcomes. To illustrate the application of these proposals we select four different outcomes of diverse nature coupled with different school climate factors for each outcome. These are: school intake, school climate factors, and teachers' actual turnover from schools (paper 1), school climate factors and teacher turnover intentions (paper 2), bullying and students' achievement (paper 3), and classroom open discussion and students' democratic attitudes (paper 4). We use secondary data with large samples of individual and schools for these purposes.

Each of these scenarios presents real challenges that serve to illustrate how these different problems in school climate research can be addressed. In paper 1, the problem of school composition and teacher turnover is illustrated. School composition is thought to be related to the outcome in an indirect way. By changing the method of how the same outcome is measured, we show that the effect of school composition present different effects. A distinctive characteristic of this paper is that it uses administrative secondary data at the national level from teachers in schools, and deals with practical challenges which are absent and less frequent in most common scenarios of large scale assessments (L. Rutkowski, Gonzalez, Joncas, & von Davier, 2010). These are data sparsity and the use of non-normal outcomes.

In paper 2, we address the problem of teachers' job satisfaction and teacher turnover intentions, and scrutinize their relationships with school climate factors. We fit a theory driven model to estimate indirect effects between these, in spite of the nested structure of the data, and illustrate how multilevel SEM models are a useful tool in this regard.

In paper 3, we use a similar approach to estimate the relationship between bullying and achievement, following indirect routes of effects via engagement variables. Here we address complexity of effects, integrating moderation of indirect effects.

Finally, in paper 4, we study the role of open classroom discussion, a school climate factor, and its relation onto the endorsement of democratic values between students, such as support for gender equality and lower perceived permissiveness towards corruption, among others. Table i2.1 serves to illustrate how well each paper addresses the school climate issues that were described earlier.

Table i2.1 Papers and school climate issues

Paper	Topic	Multi-dimensionality	Nested data	Inference level	Referent	Complexity
1.	School intake and teacher turnover	X	X			X
2.	School climate and teacher withdrawal cognitions	X	X	X	X	X
3.	Bullying and student achievement	X	X	X	X	X
4.	School climate and students' attitudes	X	X	X	X	X

Each paper receives an x where it addresses one of the issues of school climate research discussed earlier. Notwithstanding the use of school climate factors in the second study in paper 1, there were constraints on the analytic approach given the data sparsity; covariates were not partitioned into their within and between components, and instead a population estimate approach was used. That is, we used a design based method (McNeish & Stapleton, 2014, 2015; McNeish, 2014) to get valid standard errors of clustered data. Thus, it does not illustrate how measures collected at the individual level capture teachers' experience to reflect school level properties. Instead, it controls for individual experience at the population level, and centres its attention on their relations at the population level. In papers 2-4, however, analysis of between and within components of variance was possible.

Paper 1. Predicting teacher turnover from public schools in Chile, beyond socioeconomic status

Diego Carrasco
Robin Banerjee

University of Sussex, UK

Reference:

Carrasco, D., & Banerjee, R (2015). Predicting teacher turnover from public schools in Chile, beyond socioeconomic status. *Unpublished manuscript.*

Abstract

Schools with a larger intake of students from lower socio economic status are often, but not always, reported to present higher levels of teacher turnover than other schools. In two studies, we compared how this factor behaves in a survival analysis and a cross-lagged analysis approach, with attention to other potential predictive variables. In Study 1, we addressed which teachers are more at risk of leaving their first job, from which schools, and when they are more likely to leave. We used Chilean census data of teacher contracts to follow the 2007 cohort of first-time teachers through to 2013. In Study 2, we compared the turnover rates of teachers from the National Teacher Evaluation System using a logistic regression with clustered errors. Teachers of secondary school age students were at higher risk of turnover than primary school teachers and a nonlinear effect of SES was found, whereby teachers who start teaching in school with the lowest SES (81-100% vulnerable students) presented a shorter first teaching spell. Teachers are at the highest risk of turnover in their first year: 3 of 10 teachers were not working in the same school a year after starting. Crucially, Study 2 results showed that SES had only modest predictive value, in contrast to age, experience, gender, contractual status, qualifications, and teacher performance. Moreover, positive organizational features of school climate predicted a small but significant increase in the proportion of variance explained. We discuss teacher turnover as a longitudinal process, linked to schools' organizational features over time.

Keywords: teacher turnover, teacher attrition, teacher retention, teacher mobility, school workplace.

Predicting teacher turnover from public schools in Chile: Beyond socioeconomic status

A large body of research has investigated teacher attrition and turnover (for reviews see Borman & Dowling, 2008; Guarino, Santibanez, & Daley, 2006; Ingersoll, 2001; Macdonald, 1999). This work has pointed to an important connection with socio-economic status: teachers from schools serving more disadvantaged communities are more likely to leave their school during their teaching career (e.g., Borman & Dowling, 2008; Guarino et al., 2006; Ingersoll, 2001; Loeb, Darling-hammond, & Luczak, 2005). However, the explanations of the underlying individual and contextual factors remain elusive. Some explanations rely on the teacher characteristics (R. Allen, Burgess, & Mayo, 2012), others focus on the working environment of schools (Loeb et al., 2005), and still other research highlights the role played by organizational aspects of the schools (Ingersoll & May, 2012; Ingersoll, 2001).

Each perspective on its own has its limitations. Explanations which focus mainly on teacher characteristics tend to overlook the contextual and group processes that might be related to job turnover rates (Ingersoll, 2001). Having a focus on working conditions alone might ignore the heterogeneity by which different teachers and groups respond to different conditions and to relevant group processes (Dworkin, 1987). Organizational factor frameworks on the other hand, offer a compelling explanation that addresses the concerns of previous models, such as the inter-relation between individuals and environments. However, they fail to connect this with the general research on withdrawal behaviours present in the organizational behaviour literature (Griffeth, Hom, & Gaertner, 2000; Hom, Caranikas-Walker, Prussia, & Griffeth, 1992).

Consequently, the possible public policy recommendations from these different research programmes vary. When teacher characteristics are highlighted, recommendations focus on teacher recruitment and policies driving attraction of teachers with certain profiles to certain areas (Ingersoll & Smith, 2003), such as programmes to attract other professions to the teaching (e.g., Hirsch, Koppich, & Knapp, 2001). In contrast, when working conditions are emphasized, then investment in school infrastructure and salary improvements become the prominent focus (e.g., Buckley, Schneider, & Shang, 2005). Moreover, when organizational factors and management differences are taken into account, recommendations are shifted, from solely supply concerns and infrastructural worries to the necessary inclusion of school practices and the role of school principals (Ingersoll & Smith, 2003; Ingersoll, 2002).

In the present research, we use secondary public data from Chile to address the problem of turnover from public schools, with a large national sample. Our investigation provides a unique test of a very wide range of individual and school factors, encompassing all three frameworks above, in order to estimate the unique predictive value of each. The Chilean context is an interesting case for the study of teacher turnover for several reasons. Chile is one of the countries with higher economic inequality among the OECD countries, frequently at the bottom rank, with a Gini coefficient of .50 (OECD, 2014a). The large societal differences in income permeate educational differences within the population. Chile also has one of the most striking socioeconomic gradients in achievement, with a large gap between low and high SES in terms of educational achievement. It also has one of the largest opportunity gaps, the difference between low and high SES students' access to qualified teachers (Akiba et al., 2007). The latter partially occurs because more qualified teachers in Chile are allocated asymmetrically between schools in their first job: teachers with better academic qualifications are more prone to be hired in school with a higher SES intake (Meckes & Bascopé, 2012).

Beyond qualifications, however, a variety of other characteristics – both those that are attributes of the individual teacher and those that represent subjective experiences of the school context – may serve as key proximal determinants of turnover in the socioeconomically segregated Chilean society.

Teacher turnover and its relevance to school functioning

A teacher is said to attrite from a school, when she or he leaves the school where she or he teaches. Thus, attrition is necessarily a temporal phenomenon, which requires the observation of an initial time point, in which a teacher is known to be working at a given school, and a later point in time in which the teacher's presence or absence in the staff of that school can be assessed. Research has worked around two different conceptualizations of teacher attrition. The first focuses on the departure of a teacher from the teaching classroom, regardless of continuity within the same school in a different role (see for example Quartz et al., 2008), such as becoming a school head teacher or a change to another administrative role. A different conceptualization – and the one used in the present study – is the turnover approach. This approach identifies cases where an employee no longer works for a given organization after a period of time (Ingersoll, 2001), frequently a year after the initial time point (Steel & Ovalle, 1984). This definition includes both teachers who move to different schools and those who leave education entirely.

There are at least three main reasons why teacher turnover rates matter for school functioning. The most understandable reason is that employer turnover is resource costly (R. Allen, 2013). Hiring teachers is a costly task for schools, and the more frequently this happens, the more resource draining it is. For example, a study among teachers in Texas estimated that teacher recruitment costs 20 to 150 percent of the teacher's salary (Benner, 2000), whereas another study estimated the cost at 30 percent of the departing teacher's salary (Borman & Dowling, 2008). Secondly, teacher turnover has a disruptive effect on schools and a detrimental effect on pupil performance (Dolton & Newson, 2003; Ronfeldt, Loeb, & Wyckoff, 2012). Higher teacher turnover harms social relations within the schools, which alters indirectly performance and school functioning by damaging the network of social resources for the school community (Ronfeldt et al., 2012). We have already noted that teacher turnover tends to be higher in school which serves more socioeconomically disadvantaged students (R. Allen et al., 2012; Borman & Dowling, 2008), and this pattern therefore carries with it a likelihood of amplifying poor student outcomes in those schools.

Explanatory factors for teacher turnover: Individual and contextual factors

Below, we summarize the existing state of knowledge regarding a variety of individual and contextual factors thought to serve as predictors of turnover, beyond SES.

Age and experience. Meta-analysis of turnover studies show negative correlations between age and turnover (Cotton & Tuttle, 1986; Griffeth et al., 2000). The teacher attrition literature suggests that younger teachers tend to leave schools more than their older counterparts, and also that more experienced teachers tend to attrite less in contrast to their novice colleagues. However, these factors are more likely to follow a trend in form of a U shape, with higher risk at earlier stages, lower risk in the middle, and higher risk to leave when approaching the age of retirement (Borman & Dowling, 2008; Guarino et al., 2006). R. Allen and colleagues (2012) assert that age is the primary explanatory factor for the higher attrition of disadvantaged schools in England, once other market conditions are controlled for. More disadvantaged schools in England attract younger and inexperienced teachers, who are also the types of teachers who tend to leave the schools more frequently.

Gender. Females constitute the majority in the teaching profession (Guarino et al., 2006; Jensen, Sandoval-Hernández, Knoll, & Gonzalez, 2012), and meta-analytic estimates indicate higher risk of attrition for females, showing odds of 1.3 more than for males (Borman & Dowling, 2008). The increased odds of turnover for females are related to having children: Stinebrickner (2002) estimates that females are 7.72 times more likely to exit the workforce if they have a child in the current year.

Qualifications, performance, and teacher role. Teacher educational certification is considered a specific teaching skill investment, and thus is expected to predict a lower likelihood of teacher turnover (Adams, 1996; Dworkin, 1987; Feng, 2010). However, some reviewers suggest that the evidence for this effect is rather small (M. B. Allen, 2005). The picture is also uncertain with respect to higher education qualifications. Although these are often interpreted as higher investments that should buffer against burnout, quitting and intentions to leave. Yet, these qualifications may also imply a broadening of job opportunities (Adams, 1996). Therefore, higher educational attainment (i.e. highest degree) might work as protective or risk factors for quitting the school. Estimates of this factor in Chilean samples point in the latter direction: teachers with masters degrees are more prone to school turnover in contrast to their counterparts (Cabezas et al., 2011).

It is not certain whether the assessed quality of teaching performance might work as a protective or risk factor for teacher turnover. It is a specific skill certification; thus it might work as a protective factor for teacher attrition by encouraging teachers to remain in the teaching career (Dworkin, 1987). However, how this affects turnover is uncertain, because higher credentials might enhance individual employability within the teaching career. Goldhaber and Hansen (2009) followed this line of reasoning and estimate a competing risk model to teachers who stay, leave the school, leave the district, or leave the state of North Carolina in the US. The results of this research point to the second interpretation: certified teachers on high teaching performance are more mobile in comparison to their counterparts. In Chile, where teacher performance is regularly assessed through the National Teacher Evaluation System (NTES) (Manzi, Gutiérrez, & Sun, 2011), it is expected that NTES results would influence selection of new teachers and exit of low performing teachers (Taut, Santelices, Araya, & Manzi, 2010).

Teacher turnover rates are also likely to vary systematically according to the nature of the teacher's role. In the present study, we focus particularly on the phase of schooling as well as whether the teacher has a specific special needs role. Turnover is higher among middle and high school teachers (i.e., secondary school) in contrast to primary school teachers (M. B. Allen, 2005), and recent estimates from Chile reach a similar conclusions (Cabezas et al., 2011). However, both are likely to contrast favourably with special needs teachers, for whom turnover rates are expected to be higher (Billingsley, 2004; Boe, Bobbitt, & Cook, 1997).

Salary and contractual status. There is a consensus that pecuniary factors are related to turnover. Within the teaching profession, Borman and Dowling (2008) have reported an odds ratio of .66 (r=-.11)[3] to summarize the effects of studies which compare teachers' salaries with tenures of 6 or 30 years of experience. Cox proportional hazard studies of teacher attrition reaches the same conclusion: higher salaries are significantly related to longer teaching spells (Grissmer & Kirby, 1992; Imazeki, 2005; Kirby, Naftel, & Berends, 1999; Stinebrickner, 2002). In Chile, however, variation among teachers' salaries is quite low in contrast to other occupations (Mizala & Romaguera, 2000), which might imply little choice to improve income within the teaching profession. Furthermore, lower salaries are explained partially by two factors: partial contracts and low value per hour. The average contract of a teacher is for 32 weekly hours (Valenzuela, 2013) and official statistics from 2011 show that only 27.1% of teachers held contracts of at least 44 hours (MINEDUC, 2011b). Additionally, the lesser hours per contract only accounts for 11% of teachers salary gap to other professionals roles, while the lower value per hour accounts for 60% of this gap (Bellei & Valenzuela, 2010). Hence, less than a third of the teachers are in full employment (a full contract implies 44 hours of weekly work). Given this scenario, teachers' contracted hours might be a good proxy to account for differences in teachers' salaries in Chile.

Furthermore, less than 8% of the teachers work in an additional school (Valenzuela, 2013). Yet the existence of job alternatives is a known positive predictor

---

[3] An odds ratio of .66 yields a coefficient of correlation of .11 if it is transformed to d and then to an r estimate (Bonett, 2007; Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003) using dHH index transformation.

of general job turnover (Griffeth et al., 2000; Hom et al., 1992). The number of schools for which a teacher works could function as source of job alternatives; but also dealing with more than one school contract could lead to higher stress, and thus to turnover.

Organizational features. Considering only the factors discussed above would neglect a major set of contextual factors that may have a significant role to play in explaining turnover. Many authors, for example, have drawn attention to the importance of the school administration. For some authors (e.g., Boyd et al., 2011; Ladd, 2011), it is a general issue of effective leadership from the school principal, tapped by statements such as ‘the school administration is supportive and encouraging’ and ‘the school administration deals effectively with pressure from outside’. In the case of Ladd (2011), leadership is more broadly defined, including general support, trust in teachers, and the perceived appropriateness of teacher evaluations. However, interpersonal relationships within the school can be investigated more broadly. Ingersoll (2001) reports negative relations of administrative support and faculty influence with teacher turnover, as well as positive relations between student conflict and turnover. Previous research regarding accountability controls indicate that policies which make teachers responsible for their students’ academic results increase teachers’ burnout (Berryhill, Linney, & Fromewick, 2009).

General model of turnover

One limitation of the previous research on the relationship between school organizational factors and teacher turnover is the lack of a comprehensive conceptual model. Teacher turnover is certainly a multifaceted phenomenon. The broader organizational behaviour literature highlights the relevance of stressful conditions on withdrawal behaviours (Podsakoff et al., 2007) through the indirect effect these have on job satisfaction and commitment (Berry, Lelchook, & Clark, 2012; Harrison, Newman, & Roth, 2006; Hom et al., 1992). This literature emphasis the intentional and agentic nature of workers and organizational members, who under satisfactory conditions tend to remain (Kammeyer-Mueller & Wanberg, 2003), but consistently tend to withdraw under unfavourable environmental circumstances (Hanisch & Hulin, 1991). Griffeth et al.’s (2000) meta-analysis estimates present organizational commitment (r = -.27) and intentions to quit (r = .45) as the best predictors of job turnover. However, these are thought to be reflections of a variety of contextual predictors of turnover related to the working environment. In other words, negative working conditions hinder job satisfaction and commitment, which in turn lead to withdrawal cognitions, which ultimately lead to turnover (Steel & Lounsbury, 2009). This approach is consistent with the work of Ingersoll (2001) and Dworkin (1987) on teacher turnover.

A particular model of organizational behaviour research that suits the teacher context is the *job demands and resources* model (Hakanen et al., 2006; Prieto et al., 2008). In this model, there are two separate processes, which involve staff wellbeing on the one hand, and organizational commitment and teacher engagement on the other, with turnover as the ultimate outcome. The first process is a ‘strain’ process: high demands (i.e., stressors) can lead to burnout, which affects wellbeing, which in turn is negatively related to organizational outcomes. The ‘motivational’ process, in contrast, presents job resources as a boost for motivation, which in turn relates positively to organizational outcomes such as low turnover rates (Bakker & Demerouti, 2007). An interesting application of this model to teachers was employed by Bakker and colleagues (2007). The authors found support for the idea that organizational factors such as supervisor support, innovativeness, appreciation, and positive work climate were important job resources to cope with demanding student interactions (Bakker et al., 2007).

Figure 1.1. The Job Demands-Resource Model (Bakker & Demerouti, 2007)

This general model permits us to link teachers’ work environment to the larger body of research on withdrawal behaviours, and finally to a psychological and organizational consideration of teacher turnover. However, we need to situate any analysis of the organizational features encapsulated here within a broader consideration of the large number of other factors outlined earlier.

Why teacher turnover and not teacher attrition

We are especially interested in teacher turnover from schools, and not teacher attrition, because the second is a special case of the first. By convention, teacher attrition refers to teachers who leave the teaching profession. We use a turnover approach, in the most conventional way: if a teacher no longer works for given school after a period of time (Ingersoll, 2001), frequently a year after (Steel & Ovalle, 1984), turnover have occurred. This definition includes both teachers who move to different schools and those teachers who left the education profession. The classification of attrition of teachers who have left the profession can be contaminated by teachers who will return to teaching in time, outside the observed time window, thus overestimating the rates of teachers who leave the profession. Thus, in comparison, the classification of teachers as ‘no longer teaching in this school’ is a much safer categorization of events, especially if school records are the primary input to differentiate between stayers and leavers.

The following study uses the latter definition. We are focusing on teacher turnover for a given school in time (e.g., a year after) because this phenomenon has clear consequences for the schools who experience it (Kelly, 2004). Teacher turnover harms student achievement, regardless of teachers’ future career choices, because of its disruptive effects. Indeed, previous research has shown that students from classrooms with larger turnovers score lower in mathematics and language evaluations (Ronfeldt et al., 2012).

Overview of the present research

There are two popular frameworks to study teacher turnover: the cross-lagged turnover approach and the survival method approach. The first focuses on the differences between the rates of turnover, and what characteristics and factors are associated with teachers departing from a school, regularly from one year to the next. In contrast, survival analysis enables us to address jointly the questions of rates of departing and time of departure. In the present research, we use both approaches.

With the survival analysis of the first teaching spell, we addressed the questions of which teachers are more prone to turnover from their first job, from what schools, and when teachers are more likely to depart from schools. With the cross-lagged approach, we address the questions of which differences in teachers’ school experience are related to their school departure. Finally, we compare if these differences are similar between novices and experienced teachers.

Study 1: Teacher Turnover: survival analysis approach

How to measure a teaching spell?

For how long does a teacher teach in a school? One option to answer this question is to average all the tenure years of the current teachers in all the schools, which can give us the more likely duration of all the teachers currently present in the population of schools (e.g., R. Allen & Burgess, 2010). However, this would be an average estimate of all the teachers who are currently teaching, which ignores all the teachers who have left the schools and also have had a variety of different lasting tenures (Singer & Willett, 2003). Therefore, the average of teacher tenure is not an accurate estimate of teaching spells.

In contrast, a different approach is to follow a cohort of teachers in time. When a cohort of teachers is followed prospectively, some teachers will remain teaching, whereas others will leave in different years. Hence, at the end of an observation window we will have two groups of teachers: the teachers who remain at school and the teachers who left the school at different points. Thus, we can measure the amount of time in which teachers are retained in school, in contrast to the accumulated proportion of teachers who have left. These estimates permit us to build a *median life time* (Willett & Singer, 1991) which is the point in time in which 50% of teachers are still in the same school and the other 50% have left the school.

In Study 1, we used a survival discrete time analysis, to account for rates of teacher turnover from their first contract and jointly estimates when teachers have left their first school. We focus on the first school contract given the relevance that the experience of the first few years of teaching has for teacher effectiveness development (Boyd, Lankford, Loeb, & Wyckoff, 2005). Initial shorter teaching spells may retard teachers' development of professional competence. Schools that changes teachers on a recurrent basis are likely to be adversely affected, and as a whole system, higher levels of teacher turnover may harm the stock of available teachers who can reach professional maturity.

Method

Participants and data sources

We retrieved data from the annual 'Teacher Suitability Survey' (Valenzuela & Sevilla, 2013), from 2007 to 2013. This census survey registers the identity of all teachers enrolled in the school payroll. With these records, we were able to allocate all those teachers who remained at the same school after year 1, and all of those who left this school in the subsequent years until 2013. Of all those records, we selected a strict sample of teachers: only teachers with a teaching role, with a first contract in 2007, and with no contracts from 2003 to 2005, to assure these were new teachers, only from private, public and subsidized schools.

Additionally, we excluded all those cases who were not of labour active age, male=>65, and females=>60, thanks to their chances of retirement (OECD, 2013a) and we kept only those teachers who the school identified as new teachers for the educational system, with only one year or less of experience[4]. We excluded from this selection: teachers in vocational schools and teachers hired for adult education, thanks to their heterogeneity in contrast to mainstream schools. In summary, we tracked all mainstream school teachers who started a contract on 2007 over a time window of seven years, in order to create a prospective design of time to the target event, when teachers left their first school (Stinebrickner, 1998).

We retrieved the Index of Vulnerability of each school for the matching years to all the schools with the available data. By design, this index is only provided to public school and subsidized schools, not private schools. This index is used to estimate how many free schools meals should be provided by the state to the school (McEwan, 2010), and is a weighted composite score, which consists of census information from first grade and ninth grade students on health indicators (e.g., dental cavities, hearing problems, and body weight), age of school enrolment, and mother's years of schooling (Kain, Uauy, & Taibo, 2002). This index follows a negative correlation with family income, and is often used as a measure of school socioeconomic differences (Agencia de Calidad de la Educación, 2012b).

[4] The question of how many years a teacher has in the school system is considered ambiguous. By year 1, a teacher in continuous time would have just a few months of experience, but in common language it is often the case to say that a given individual is in year 1 once he or she has started the academic year.

The inclusion of this covariate led us to fit two different models: one to compare the whole population of first teaching spell between public, private and subsidized schools, ignoring socio economic status students (SES) intake; and a different model for publicly funded schools while accounting for the SES intake of the school.

Our final sample consisted of 7,458 teachers (Females 75.8%, mean age 28.66 years, STD=6.72), from schools in different regions of the country (Metropolitan Area=40.33%, Region V=11.21%, Region VIII=11.20%, and remainder=37.26%), and from different type of administrations (public=22.37%, subsidized=65.73% and private=11.91%).

Predictor measures

SES of the school intake. Chilean educational policy use weighted index to assess students in terms of vulnerability and economic disadvantage (Kain et al., 2002). Those schools with a higher proportion of students who meet this criterion receive differential funding and free school meals, among others forms of support. This index varies from 0% to 100%, indicating the proportion of students who are entitled to benefit. In our models, this index was discretized into four dummy categories, in order to deal with the lack of linearity relation of this predictor to the outcome. We follow Singer & Willet (1993), inferential and exploratory methods to reach a sensible set of dummies, to account in the best way the non-linear relation between SES school intake and teacher’s trajectories. These categories were 0% to 50% IVE Students (reference category), 51% to 70% IVE Students, 71% to 80% IVE Students, and 81% to 100% IVE students.

Rural. A dummy variable was introduced to distinguish schools from urban areas and cities, and schools from less populated areas, such as rural villages.

School size. We retrieved the number of teachers per school in 2007.

Class size. We retrieved the number of enrolled students per school, and the number of grades (age groups), and estimated the average number of students per class. We discretized this index, to account for its non-linear effect. We created three category dummies: school with fewer than 25 students per class, between 25 and 33 students and schools with 34 students per class or more.

Age. Teachers’ age was computed in years, from their date of birth to March of each year, which is the beginning of the academic year in Chile. This variable was discretized into three groups ranges: young teachers (younger than 30 years), mature (between 30 and 40 years), and senior (more than 40 years). The middle category was used as reference, thus not included in the analysis.

Gender. Teacher's sex was computed as a dummy variable, in which Male is the reference category (with a zero value), and Females are identified with value of 1.

Hours by Contract. We used hours by contract as a proxy for possible differences in salaries between teachers. This variable varies from 1 to 65 hours per week. This variable was discretized into 3 ranges: teachers with fewer than 29 hours a week, those with 30 to 35 hours a week, and those with 36 or more hours a week.

Type of Contract. We created a dummy variable to account for those teacher who hold an indefinite contract.

In how many schools the teacher is working. Teachers can be hired to work for more than one school: some teachers can hold partial contracts in several school to reach higher salaries and full employment, rather than holding a full contract of 44 hours a week for a single school.

Teaching role. Teachers in Chile are broadly classified as Nursery, Primary (students ages groups 5 to 13), Secondary (students ages groups 13 to 17), and Special Needs Teachers. These four different teaching roles contribute to significant heterogeneity within the teaching profession: they express differences in job design (daily activities), and also differences in curriculum. We created dummy coded variables for this factor, one for each teaching role, with Primary school teachers as the reference category.

Table 1.1. Study 1 Sample Descriptives

Analysis plan

We first estimated the odds ratio for each factor predicting teacher turnover, without adjusting for any of the other covariates. During this first step we checked for non-linear effects among continuous covariates, via graphical and inferential methods (Singer & Willett, 1993). We then turned these into categorical variables to express their relations to the outcomes. Then, we estimated the unique predictive value of each factor on teacher turnover including all selected covariates. We fit two models, one including all schools, and one including only the public and subsidized schools.

We used a discrete-time survival model with proportional odds ratio (Willett & Singer, 2004), within a latent variable framework (Masyn, 2003; Muthen & Masyn, 2005). Survival models or Event history models enable us to estimate whether teachers are retained or "survive", or if teachers have left the school, as well as when the latter event has occurred (Willett & Singer, 1991). In these models, time is recorded as discrete periods, given the assumption that attrition and turnover from schools is more likely to occur at the end of the school year (Quartz et al., 2008). We also corrected our standard errors with a sandwich estimator to account for possible dependency of teachers nested within the same school (Muthen & Masyn, 2005). This strategy was deemed more appropriate than a variance component model (e.g., multilevel) given the high proportion of singletons in our sample (McNeish, 2014): 27.1% of our observed cases had no other colleagues in the dataset from the same school.

To check for basic assumptions of the fitted models, we compared our selected model to other alternative specifications: a model with non-parametric frailty, which accounts for unobserved heterogeneity (Masyn, 2003), and a model with non-proportional odds, in which covariates are allowed to have time variant effects (Willett & Singer, 2004). However, the comparison between the estimated Bayesian Information Criterion (BIC) of each model specification favoured our chosen model specification (see Appendix (Paper 1), Table 1.11).

Results

We assessed the model fit via a Likelihood ratio test (Muthen & Masyn, 2005). This Log likelihood ratio Test, is corrected using the scaling correction factor produced by the MLR estimator in MPLUS (Masyn, 2014). Our tested models present a better fit when compared to their null counterpart (Model 1: LRT = 244.43, df = 16, $p < .01$; Model 2: LRT = 246.18, df = 18, $p < .01$).

Table 1.2. Fit Statistics of survival Estimates

Table 1.3. Survival Estimates of duration of first teaching job

----

In Table 1.3, estimates of the two survival models are provided. In the first set of columns is the model in which all teachers from all schools are compared, whereas in the last set of columns, we display the estimates for the state funded schools only. In general, positive coefficients indicate higher risk of turnover. If we look at the School differences block of predictors, the more salient effect in Model 1 is the difference between teachers who start working in Private schools to the rest of the teachers. Teachers in the private sector are 1.16 times more likely (b=.15, SE=.08, $p<.05$, *hOR*=1.16) than Public teachers to leave their first school contract, in any given year. In contrast, when only the publicly funded schools are compared, taking into account the SES intake, the most vulnerable schools category is the most prominent predictor for shorter first teaching spell (b=.25, SE=.07, $p<.01$, *hOR*=1.28). For this cohort, teachers working in the most vulnerable schools are 1.28 times more likely to turnover, in contrast to schools with Higher SES (between 0% and 50% of IVE student's intake).

Figure 1.2: Fitted survivor and hazard probability describing the risk of leaving the first school contract, by SES intake and Teacher Role, Contract Status and Hours per week

----

When the median life time (Singer & Willett, 2003) of these two groups is estimated this differed by about a year. Teachers in schools with a High SES intake (0-50% IVE students) are expected to have closer to an additional year of teaching (median lifetime = 3.1 years) on their first teaching job, in comparison with teachers in schools with the Lowest SES intake (81-100% IVE students; median lifetime = 2.2 years). These estimates can be identified graphically by looking at the time scale of each line, in each plot from Figure 1.2, which crosses the horizontal line of the 50% Survival probability.

When the rest of the covariates are scrutinized, estimates do not differ between the models and their confidence intervals overlap greatly. Thus, the estimates of the Model 1, which describes the relationship of the factor in all the schools, will be used to describe the overall results for the rest of the factors.

When life factors are taken into account, age of entry and sex are significantly associated with teachers' first spell duration. Those teachers who start their first teaching spell when they are younger than 30 years of age, are more at risk (b=.12, SE=.07, p<.01, hOR=1.13) of leaving their school, than their counterparts of 30 to 40 years. Males have shorter tenure on their first contract in contrast to Females: Male teachers are 1.11 times more likely to leave their first job each year (females, b=-.11, SE=.04, p<.01, hOR=.90; males hOR=1.11).

Working conditions also play an important role. Teachers who hold an indefinite contract are at lesser risk than their counterparts are (b=-.22, SE=.04, p<.00, hOR=.80). Teachers with different contracts are 1.25 more likely to leave than teachers hired with an indefinite contract. Similarly, teachers with less than the mode of contracted hours (30 hours a week), are 1.32 times more likely to leave than teachers who are hired for 30 to 35 hours a week.

Finally, between the different teachers in the mainstream schools, Secondary School teachers are at higher risk of turnover than the reference category of Primary School teachers, when all the previous factors are controlled (see Figure 1.2). Secondary School teachers are 1.38 times more likely to turnover from their first contract than Primary school teachers (b=.32, SE=.04, p<.01, hOR=1.38). If we estimate the median lifetime using the parameters from Model 1, Secondary School teachers displayed an expected median lifetime of 1.8 years, whereas Primary School Teachers have a median lifetime of 2.6 years, and similarly Nursery Educators present median lifetime of 2.8 years, and Special Educators a median of 2.3 years. Overall, Secondary School teachers are expected to have shorter first teaching spells.

Study 1: Conclusion

We can now review the main effects which relate to our questions of Who, When, and from Where. **Who?** Among the different population of teachers, secondary school teachers presents the shorter teaching spells (Median Life time of 1.8 years). By the start of year 3 (interval from year 2 to 3 in Figure 1.2), only 46% of those who started their first contract in year 2007, remain in same school. **When?** Teachers are at a higher risk in year 1, where 3 out 10 teachers will leave their first teaching contract. Hazard probability plots in Figure 1.2, peak at year 1. **From Where?** Teachers have shorter first teaching spells when they are working in schools from the lowest SES range (80% to 100% of IVE Students intake). However, given this model is a proportional odds ratio model, the effects are additive. Thus, we can estimate a highest risk profile: teachers from secondary grades, males, with less than 30 years of age at entry, working in a subsidized school, with the lowest SES intake, in their first year in the job. Among all, this teacher profile is at the highest risk of moving away from that school.

These results motivate two questions: what other school differences may explain different rates of turnover, and additionally, are these differences similar between novice and non-novice teachers? With the survival model, we have followed novice teachers. Thus, with these results, we do not know what explains turnover from schools for teachers in general. In Study 2, we address these two questions.

Study 2: Teacher turnover: cross lagged approach

Method

Participants and data sources

We used the 2010 cohort from the National Teaching Evaluation System (Santelices & Taut, 2011; Taut et al., 2012), comprising a total of 11,061 public school teachers. To complement this survey, we cross-referenced this information with the annual ‘Teacher Suitability Survey’ (Valenzuela & Sevilla, 2013) which records the number and identity of teachers enrolled in school payroll, leaving us with a sample of 10,895 cases.

With these records, we were able to identify all those teachers who remained at the same school, and all of those who left a year after. Additionally, we collected school information from different Ministry of Education records such as indexes of school size, school socioeconomic intake, and school’s academic results.

Our final sample consisted of 10,895 teachers (Females 69.4%, mean age 44.97 years, SD=11.45) from different regions of the country (Metropolitan Area 28.3 %, Region V 10.5%, Region VIII 9%, and remainder 52.2%)[5].

Our key dependent variable is a binary indicator of teacher turnover. We classified each teacher from the 2010 NTES cohort as a stayer (0) or as a leaver (1). A teacher was considered a stayer if she or he appeared to be working in the same school in year 2011. A teacher was considered a school leaver if she or he appeared in a different school or was absent from the Teacher Suitability Records for the year 2011. In total, 14.02% of teachers were classified as leavers.

The majority of our predictor variables were derived from the 2010 NTES complementary survey conducted by the Centre of Studies, Ministry of Education, Chile (Taut et al., 2010). This is a paper and pencil survey completed on a voluntary basis by every teacher selected for evaluation in the given year. In this, teachers give their informed consent to release their responses for research. It covers topics to characterize teachers' qualifications, their work experience, and finally their opinions and experience regarding the process of evaluation.

Predictor measures

Socioeconomic status of the school intake. At the start of the academic year (March), schools are surveyed by the 'Junta Nacional de Auxilio y Becas' (JUNAEB) which estimates the proportion of students under vulnerability conditions and determines the amount of free school meals for each public and subsidized school (JUNAEB, 2005). This index is a proportion, which varies from 0 to 1, and is a fix index per school. We averaged the primary and secondary level index, to characterize the school as whole. Afterwards, this index was discretized into 5 categories (see Table 2).

Rural. A dummy variable was introduced to distinguish schools from urban areas and cities, and schools from less populated areas, such as rural villages. School from non-urban areas were flagged with a value of 1.

School size. We retrieved the number of teachers per school in 2010.

[5] These are raw estimates, with no clustering correction.

Class size. We estimated an average class size per school (total enrolment divided by total number of classes). This index was further discretized into three dummy variables (see Table 1.4).

Academic results. School results were retrieved from the 'Sistema de Medición de la Calidad de la Educación' (SIMCE). These are academic results evaluated by a standardized test, with a census design. Each year, students from grades 4, and alternatively grades 8 and 10 are evaluated in Math, Language and other topics (Agencia de Calidad de la Educación, 2012a). Math scores from each schools were retrieved for 2010, standardized and average, for this study.

Age. Teachers' age was computed in years, from their date of birth to March of the year 2010, which is the beginning of the academic year in Chile. This variable was discretised into three dummy variables, to account for the non-linear relationship this factor is thought to have with turnover. Ranges were for young teachers (less than 30 years), mature (between 30 and 40 years), and senior (more than 40 years). The middle category was used as reference, thus not included in the analysis.

Years of experience. In the NTES survey, teachers state for how many years they have been working in their teaching career. Again, to account for the expected non-linear relationship, the variable was discretized into four dummies, for ranges of 0 to 3, 4-9 years, 10-20 years, and more than 20 years. This final range was used as a reference in the analysis.

Gender. Teacher's sex was computed as a dummy variable, in which Male is the reference category (Female=1, Male=0).

Type of Contract. We created a dummy variable to account for those teachers who hold an indefinite contract.

Hours by Contract. As explained earlier, we used hours by contract as a proxy for possible differences in salaries between teachers.

In how many schools the teacher is working. We computed a count variable for how many schools a teacher works for, during 2010.

Workload. Teachers indicate, with a Likert type item of 5 levels, if they feel they had enough time to complete their work: 'The amount of tasks I need to do at this school exceeds the time I have'.

Teacher performance. We included a dummy variable identifying teachers who recorded a satisfactory or outstanding result in the NTES evaluation (the remainder were coded as 0).

Teaching role. Teachers in Chile are broadly classified as nursery, elementary (ages 5 to 9), middle (ages 10 to 13), or secondary (ages 13 to 17) teachers, who are all differentiated from special needs teachers. These five different teaching roles comprise high heterogeneity within the teaching profession: they express differences in job design (daily activities) and also differences in curriculum. We created dummy coded variables for this factor, one for each teaching role. We left elementary and middle teachers as reference categories to be consistent with study 1.

Teachers' qualifications and training. We included additional dummy covariates for teachers' further training: PhD, Master degrees, Post Titles, Diplomas, Training above 60 hours, and Training above 60 hours specifically within the teaching subject.

Organizational factors: **Principal leadership**. Teachers indicated if the school head teacher or school principal supports them in case of difficulties, if the school management team uses effective procedures to evaluate them, and if the school holds a shared vision of goals. These 3 items were responded to using a Likert type scale of agreement (5 levels), and then were mean scored at the teacher level. This mean score yields a Cronbach's alpha of .83. These particular measures of perceived leadership emphasize the supportive character of the school principal role (Cerit, 2009).

Positive interpersonal relations. Teachers gave their level of agreement if they thought all teachers feel respected in the school community, and indicated if they thought students and teachers got along, using a Likert type scale of five levels. These two items were mean scored at the teacher level (α = .75).

School priorities. Schools can prioritize academic results and practices, as well as the relational aspect of their school community (Hoy, 2012; Swars, Meyers, Mays, & Lack, 2009). In the NTES survey, teachers responded how important the following aspects were for the school management: academic outcomes (e.g., the teacher evaluation results, the SIMCE results, and the rate of student approval) and the relationships within the school community (the relationships of the teacher with their students, and the relationship between colleagues and the directive school staff). All these aspects were rated on a Likert type scale, indicating the level of relevance (1 for least relevance up to 5 for the highest relevance). Each group of items were mean scored separately to create two different scores. The first accounts for higher school priorities for academic results (α = .73) and the second accounts for higher school priorities for interpersonal relations (α = .78).

All Likert type scales present an acceptable factor configuration, when a confirmatory factor analysis is fitted (CFI=.96, TLI=.93, RMSEA=.056, SRMR=.03). Samples of these items are included in Appendix (Paper 1).

Prior intention to move schools. With a Likert type scale (from 1 to 5) teachers responded if they wanted to remain in the same school or if they wanted to work in a different school (higher score reflects a greater intention to move to a different school).

Analysis plan

First, we estimated the odds ratio for each factor predicting teacher turnover, without adjusting for any of the other covariates. Then, we estimated the unique predictive value of each factor on teacher turnover.

We avoided using a discrete time survival model to approach this question because of the left censoring and left truncation of this sample of teachers. These two cohorts of teachers are left censored because their teaching careers started at different time points and we cannot easily retrieve reliable records of when their first contract started; in fact, data from the Teacher Suitability Survey are considered reliable from 2004 onwards (Valenzuela, 2013). The sample is also left truncated, because some of the teachers have already presented some level of migration between schools (i.e., their current position is not their first school job) and therefore a portion of this sample could consist of long survivors for the events of leaving a school job or leaving education altogether for a period of time. Therefore, a binary logistic model with clustered errors was preferred, in order to estimate the relative relation between different factors and turnover.

Teachers are nested within schools, and when observations are clustered, the most recommended modelling approach is the use of multilevel models. However, these models require certain conditions to be met to guarantee the parameter estimates are reliable enough. Common desirable data features are: to have at least 50 clusters, and at least 5 observations per cluster (Maas & Hox, 2005). A scenario in which these models suffer greatly from bias is when there is high data sparseness, which is characterized by a high proportion of observations with too few cases per cluster, and/or a high proportion of singletons (one case per cluster). Multilevel models under high sparseness tend to overestimate the variance components, which can cause problems with the fixed effects and consequently with any inferential test; this is specially the case with binary outcomes (Clarke, 2008; McNeish, 2014).

Given the NTES survey design, the present study includes a sparse sample of teachers in each school. To be selected for evaluation, teachers must work in a public school and have at least one year of experience (Manzi et al., 2011); if the teacher gains an unsatisfactory result, it would be evaluated again in the following year, but if the teacher gains a satisfactory result, evaluation may not take place for another 4 years. Given this design, the number of teachers who are evaluated for a given year in a given school varies widely, showing data sparsity. For this cohort, around 30% of the schools are represented by singletons and only 25% of the surveyed teachers are allocated in schools with at least 5 teachers per school[6].

Although it is not settled if multilevel or clustered errors are preferable for all nested data structures (see Bell, Morgan, Kromrey, & Ferron, 2010; Clarke, 2008; Gelman, 2007; Sauzet, Wright, Marston, Brocklehurst, & Peacock, 2013), we opted for the latter given this method tolerate data sparseness (Clarke, 2008; McNeish, 2014) and require fewer assumptions (Primo, Jacobsmeier, & Milyo, 2007). Multilevel model of categorical data do not often meet the normality assumption of level 2 error distribution, thus biasing the parameter estimates (Wright, 1997). When assumptions are not met, robust standard errors seems a reliable choice (Maas & Hox, 2004) and yield estimates that do account for the nested structure of the observations (Antonakis et al., 2010) though without modelling the between-cluster variation, in contrast to multilevel models (Gelman, 2006).

When observations are clustered, single level logistic regression and linear regression estimates tend to bias downwards the standard errors, even when there are as many as two observations per cluster (Clarke, 2008). Unlike simple logistic regression, design based methods (e.g., generalized estimated equations, sandwich/robust standard errors) allow us to account for the data cluster dependency of the observations, correcting the standard errors of the models, thus protecting from Type I error, when clustering is ignored. These methods provide population averaged regression coefficients, variance components are not biased under high sparseness, and estimates are interpretable in the same way as single level regression coefficients, especially for discrete outcomes (McNeish, 2014). For all these reasons, logistic models with

[6] In the final model with all covariates 38.06% of teachers are clustered in schools with at least 5 other teachers present in the study.

sandwich estimator, specifying schools as the cluster of interest, were fitted using MPLUS v7 (L. K. Muthén & Muthén, 2012), to account for clustering and deal with data sparsity.

Table 1.4. Study 2 Sample Descriptive Estimates

Table 1.5. Study 2 Sample Descriptive Estimates (… continuation)

Our primary analysis focused on the unique contributions of the various predictors. All variables were entered into a logistic regression model in blocks, starting with overall school difference as indicated by SES intake, rural location, school size, class size, and academic results. Afterwards, we included a block of life course factors, including age (estimating all but the middle category for 30-40 years), teaching experience (estimating all but the Experienced teacher category, with 4-9 years of experience) and sex. Then we included differences in terms of contract, hours by contract and for how many schools the teacher works, and a measure of workload, to account for different working conditions. Afterwards, we added dummy variables for teacher qualifications, role, and the NTES result for that year. Next, in order to obtain a conservative estimate of the significance of organizational features in predicting turnover during the subsequent year, we controlled for pre-existing intentions to move school during 2010, and then we included the organizational factors: school leadership, positive school relations and the perceived prioritization of academic and relational issues. All models were fitted using MPLUS v7 (L. K. Muthén & Muthén, 2012), with MLR estimator with clustered errors, including dummy variables and grand mean centred variables for non-binary covariates. Given some covariates are at the teacher levels and others at the school level, in order to account for the nesting of observations a sandwich estimator was employed with the school as a cluster (Muthen & Masyn, 2005; B. Muthén et al., 2002).

Results

Table 1.4 provides descriptive statistics for all variables in the analysis, including means, standard error of the mean, the hazard odds ratio for teacher turnover for each covariate and the number of valid cases per covariate. All the estimates were calculated using sandwich estimator, similarly to the logistic models. Most of the included factors show some predictive value.

The final model (model 6 in Tables 1.7 and 1.8), which includes all the covariates, was found to explain 21.5% of observed variance in turnover. This was tested against a null model in which all slopes are fixed to zero. Using the corrected Likelihood Ratio Test (Asparouhov & Muthén, 2005), results favour our final model ($LRT(33) = 558.52$, $p < .00$) in contrast to the null model. Tables 1.7 and 1.8 shows the estimates of the fitted models for each selected block of variables. In this table all estimates are unstandardized, and for the final model an additional column was added to display the hazard odds ratio for teacher turnover, which corresponds to the exponentiated version of the estimate. In general, an odds ratio greater than 1 indicates an increase in the likelihood for a teacher to turnover from the school. Estimates for the dummy variables consist of an increase relative to the excluded category (i.e. reference category), whereas for continuous predictors, estimates reflect a one-unit increase compared to the grand mean. Variables with hazard odds ratio lower than 1 have a positive relation to teacher retention within a one-year period, and a negative relation to teacher turnover.

A panoramic comparison of all the models, relying on the Pseudo R^2 (McKelvey & Zavoina, 1975), shows that the life course factors are the most salient block. This set of factors accounted for an incremental explained variance of 9.7%, after controlling for school differences such as SES intake, class sizes, location and size of the workforce. Contractual status, number of contracted hours, number of schools, and workload added a further 2.9%, with teacher qualifications and roles adding 3.4%, prior intentions to move school adding 2.9%, and organizational features a further 1.3% of explained variance. To be cautious regarding the plausible multi-collinearity of the selected factors, we estimate the Variance Inflated factor of all included variables (see Table 1.5). All of these were lower than 10, which is common rule of thumb for concern (O’Brien, 2007).

In Model 1, Rural Schools and schools with higher academic achievement were found to be associated with a lower likelihood of turnover. However, after age, experience and sex are entered into the model, only Rural Schools appears significant from the first set of school factors (*b*=-.24, SE=.12, p<.05). Thus, from Model 2, estimates show that teachers from urban schools are 1.27 times more likely to turnover than teachers from rural schools.

As expected, Model 2 results showed that more experienced teachers are less likely to turnover, in contrast to teachers with less experience (less than 10 years). Senior and Veteran teachers have adjusted Odds Ratios of .65 (*b*=-.43, SE=.10, p<.01) and .34 (*b*=-1.08, SE=.12, p<.01) thus implying that the reference category, Teachers with 4 to 9 years of experience, have higher chances to Turnover (*hOR*=1.54).

When working conditions are included, in Model 3, the novice teacher factor from the previous model ceases to display a significant result. However, older and more experienced teachers are still at lower risk of turnover. The most salient aspects from this block relates to the teachers' contract: teachers with an indefinite contract (*b*=-.59, SE=.09, p<.01, *hOR* =.56) and teachers with higher contracted hours (*b*=-.03, SE=.01, p<.01, *hOR* =.97) are less likely to leave the school in the subsequent year.

In Model 4, a full block of teacher qualifications and roles was added to the model. A positive NTES evaluation is related to teacher retention: teachers with higher results in the evaluation are more likely to remain in their school (*b*=-.52, SE=.07, p<.01, *hOR* = .59). In terms of role at school, the results indicate that teachers from the special education field are much more likely to turnover than primary teachers (*b*=.72, SE=.13, p<.01, *hOR* = 2.05). The relationship between further training and turnover shows a different pattern. Teachers with Master degrees are more prone to teacher turnover than their colleagues with no further training (*b*=.50, SE=.12, p<.01, *hOR* = 1.65). The reverse effect is observed for more specific skills training: teachers with post titles and trainings above 60 hours are less likely to turnover (*b*=-.22, SE=.08, p<.01, *hOR* = .80).

Table 1.7. Logistic regression analysis of the likelihood of teacher turnover

Table 1.8. Logistic regression analysis of the likelihood of teacher turnover

(continuation)

After conservatively controlling for any pre-existing intentions to move school in 2010, organizational factors were found in Model 6 to explain a small but significant degree of variance in turnover. From this block, perceptions of school leadership (*b*=-.15, SE=.06, p<.01, *hOR* = .86) and prioritization of relational matters (*b*=-.15, SE=.06, p<.05, *hOR* = .86) showed negative relations to teacher turnover, thus playing a protective factor role to the risk of turnover.

In a further step, we divided our sample between novice and non-novice teachers, to check if these relationships were stronger or weaker between these two groups. We created two groups, those with 3 years or less years of experience (novice = 22.1% of the sample, N=1460), and the remainder with 4 or more years of experience (non-novice = 77.9%, N=5145). We fit these models in a multigroup setting, to compare a model in which all the estimated parameters are held equal between both groups, in contrast to the model in which all parameters are allow to vary between the two groups. The model comparison via a Wald test, favours the latter specification (Wald (30)=109.38, p<.01).

Table 1.9 depicts these results in detail, and results are different for each generated group. For example, working conditions appear to matter more for non-novice teachers. Holding an indefinite contract also explains more difference between teachers with more years of teaching (*b*=-.93, SE=.12, p<.01, *hOR* = .39, CI 95% [.31; .50]) and teachers with less experience (*b*=-.08, SE=.18, p=.64, *hOR* = .92, CI 95% [.65; 1.30]). Yet, contracted hours have a similar relationship for both groups. Intentions to move to a different school are only significant for the more senior group (*b*=.17, SE=.03, p<.01, *hOR* = 1.19, CI 95% [1.12; 1.26]). When organizational factors are inspected, there is a different pattern of results. For the novice teachers there is a negative relationship of school leadership and turnover (*b*=-.24, SE=.10, p<.05, *hOR*=.79, CI 95% [.65; .95]). There is no relationship with positive school relations, but there is a positive relation with Academic Monitoring (*b*=.26, SE=.12, p<.05, *hOR*=1.29, CI 95% [1.01; 1.65]); and finally there is a negative relation with Staff relations as a priority for the school (b=-.21, SE=.10, p<.05, *hOR*=.81, CI 95% [.66; .98]). In the case of the more senior group, only Positive School relations explain different rates in which teachers from this group turnover (*b*=-.14, SE=.07, $p<.05$, *hOR*=.87, CI 95% [.77; 1.00]).

Table 1.9. Logistic regression estimates, by groups of Novice and Non-Novice teachers

----

Study 2: Conclusion

Over and above School SES intake and teacher characteristics, working conditions and school climate factors account significantly for different rates of teacher turnover. In schools in which the head teacher and his or her team prioritize the relationships with their teachers to a lesser extent and where there are lower levels of school leadership (less supportive, less fair and with a lack of shared vision among colleagues), teachers are more likely to leave. These results are partially moderated by teachers' experience. First, novice teachers (<4 years of teaching) are more sensitive to accountability pressures. Novice teachers in school with higher levels of academic monitoring are more prone to turnover, whereas they appear to respond positively to a supportive school leadership and a management that cares about the interpersonal relationships between members of the school community. In contrast, more experienced teachers are more responsive to direct interpersonal relationships between colleagues and students. More experienced teachers are thus more likely to leave schools in which they do not feel respected and in which they feel teacher and students do not get along.

General Discussion

The present results corroborate many aspects of the existing literature, using robust tests on a wide array of covariates, in data from a large national sample. These also extend previous evidence by identifying organizational aspects of the school as a worthy factor to broaden the scope of the study of risk factors for teacher turnover. Life course factors, which capture variations in the profile of the workforce, are of greatest significance, substantially increasing the proportion of explained variance in comparison with school differences such as low SES and school Academic Achievement, which in fact is rendered non-significant after age, experience, and gender are included (Study 2). However, this was not the case, when first teaching spell was the focus of analysis (Study 1), whereby teachers with Lowest School SES intake, are 1.28 more likely to turnover, from their first teaching contract. This contrast highlights the relevance of conceptualizing teacher turnover as a longitudinal process.

Teachers' contractual status and roles are also relevant to turnover. These factors explained difference in rates of turnover on both approaches: to explain first teaching spell, and to explain rates of teacher turnover from one year to the next. Notably, teachers' NTES evaluation results have a positive effect on teacher retention, unlike results from previous research on teacher performance certification (Goldhaber & Hansen, 2009).

Results regarding teachers' qualification are consistent with previous studies on the same population (Cabezas et al., 2011). More qualified teachers (teachers with a Master Degree), are more likely to turnover than their peers. Yet, we did not replicate the same effects with teachers with PhD degrees, for whom the previous studies also found positive estimate to turnover. Additionally, we included dummies for specific job training (Training >60 hours). Teachers with more than 60 hours of job training are less likely to turnover, in contrast to their colleagues. This overall pattern of results is consistent with human capital theory; it is expected that teachers with more specific training, i.e. more investment, would show less turnover (Dworkin, 1987). However, Dworkin (1987) explains this may be an issue of entrapment: workers with less translatable skills to other occupations might not be able to move from their current positions, due to their lack of access to other choices. The opposite occurs when workers do have more translatable qualifications.

Previous research in Chile on teachers' career trajectories coincides with the present work with respect to the role of school climate factors (Bravo, Urrutia, & Peirano, 2006; Cabezas et al., 2011; Valenzuela & Sevilla, 2013). In Study 2, we show that – notwithstanding the relatively small proportion of variance explained – teachers' subjective experiences of organizational features do predict different likelihoods of leaving their schools within the following year, even after adding a strict control for any pre-existing intention to move (which of course may well have reflected organizational features too). It is particularly interesting to note that along with supportive leadership, a perceived prioritization of relational matters is a protective factor against turnover, in contrast to the opposite effect for a prioritization of academic matters.

Furthermore, these results were heterogeneous between novice and non-novice teachers. Novice teachers were more sensitive to difference in school climate factors (leadership, management of staff relations, and academic monitoring). These factors seem relevant for schools to take into account, for example, in staff induction and retention during the first years.

General models of school effectiveness highlight the relevance of schools' emphasis on academic success in order to drive improvements (Scheerens et al., 2003). However, a high monitoring of academic results might work as a job demand, thus acting as a straining rather than motivating factor for teachers (cf. Bakker & Demerouti, 2007; Dworkin, 2001); thus a single school climate factor from school effectiveness research can display positive and negative relations to different school outcomes.

In contrast, supportive leadership and a perceived school *commitment* to enhancing relational matters (perhaps more than the actual experience of positive relationships) may be serving as job resources that promote positive teacher engagement (Bakker & Bal, 2010; Bakker et al., 2007) and reduce the risk of turnover (see also Allensworth, Ponisciak, & Mazzeo, 2009; Swars et al., 2009). As some theoretical models of school effectiveness combine academic emphasis and a focus on positive relationships (Hoy, 2012), these marked differences in correlates with turnover could easily be masked. The present study shows that separating factors that function as resources for coping and those that work as demands and challenges gives a richer picture for the comprehension of teacher turnover.

Limitations and Further Research

Notwithstanding the large sample size from across Chile, one of the main limitations of these studies is the possibility of sample bias. Cohorts for the NTES evaluation vary from year to year in composition, and our sample is further restricted by the schools/age groups who participate in the SIMCE academic test. However, the inclusion of different covariates to control for the heterogeneity of teachers may help to reduce Type II errors by reducing error variance (Huck, 2012). We also acknowledge that the estimates of SES intake might be underestimated, because we do not have the continuous measures of students' disadvantage, and are relying instead on the discretized version of this index in each school, which was designed for a different purpose (Kain et al., 2002). Nonetheless, the results do point to the need to go beyond such characteristics in order to appreciate the individual and contextual factors that could explain significant variance in teacher turnover.

A further limitation of the study is that there is a temporal misalignment between the ongoing process of teacher arrival and adjustment to the school, and the unfolding process of teachers deciding to stay or leave (T. H. Lee, 2013). In Study 2, we did our best to account for life course factors such as age, experience and gender to account for these issues, but a stronger test would follow teachers over a larger timescale from the first point of holding a contract at a school. In Study 1, we fit a survival model to describe the relationship between school differences, life course factors and working conditions to first teaching spell. We subjected our fitted model to sensitivity analysis to check that our estimates were robust to different model specifications, including unobserved heterogeneity (see Appendix (Paper 1)). Our estimates appear trustworthy in light of the different specifications. However, these estimates cannot account for the relationships between teachers' school experience and rates of turnover that we showed in Study 2.

A more complex design should include measures of school climate factors, integrating both approaches, accounting for the survival nature of the outcome, its longitudinal process and clustered structure of the data (teachers within schools). The proposed latent framework for discrete time survival used in Study 1 can be expanded for multilevel and longitudinal data structures (Muthen & Masyn, 2005; Petras et al., 2011). However, the consequential attrition of the data is less straightforward to deal with, which may discourage research in this line. Current advances in missing data analysis promise progress in this regard for studies aiming to integrate time variant effects (longitudinal analysis) with survival outcomes (Enders, 2011; Sterba, 2015).

A final limitation of our secondary analysis relates to the restrictions that the original data collection design imposes on us as researchers using the data (Rogers, Anderson, Klinger, & Dawber, 2006). It is possible that the strength of the relations will appear weaker because of the relatively small number of items available to measure each included construct. Although our multi-item measures did show good internal consistency and good overall fit in confirmatory factor analysis, further research using more elaborate scales capturing organizational features of school climate factors would be helpful for advancing our understanding of the role played by such characteristics. A good source for secondary data analysis, is Teaching and Learning International Survey (OECD, 2014e), which surveys a random sample of teachers within random samples of schools in more than 30 countries.

In future, more detail is needed to evaluate competing risk models for differentiating between those who move schools and those who choose to leave the teaching profession altogether (see Kukla-Acevedo, 2009). There are effects which may appear relevant under survival approaches, but tend to be underestimated under binary logistic models (T. H. Lee, 2013). Again, there is a need to jointly address questions of "whether" and "when" teachers leave schools, for truly understanding the process of turnover over the timescale of a teacher's career. Teacher turnover is a longitudinal process with time variant and invariant covariates (Murnane, Singer, & Willett, 1988; Singer & Willett, 1993; Willett & Singer, 1989, 1991). Cross sectional designs restrict models to only time-invariant effects, whereas in terms of the unfolding process several factors change. For example, working hours and contract type vary over time, and those changes may influence the outcome of teachers staying or moving from the schools. Moreover, teacher, principal and student mobility might interrelate (Plecki, Elfers, Loeb, Zahir, & Knapp, 2005; Swars et al., 2009). A full assessment of these processes, alongside changing subjective perceptions of organizational features, requires a planned longitudinal design for prospectively studying teacher turnover. Study 1 is an input in this direction. Its estimates provide a guideline regarding what is the hazard baseline and shape for first time teachers. In order to capture the median lifetime of the first teaching spells, we now know a longitudinal study should aim for at least 3 to 4 years of follow-up.

There are at least three other factors related to teacher turnover rates that relate to much broader societal issues, which should be considered when planning future studies. First, population growth (or lack thereof) is related to the number of student enrolments and consequently teachers' job demand. For example, the baby boomers phenomenon during the 1960s and early 1970s in the US had an effect of 25% increased school enrolment and the demand for entry level teachers increased dramatically (Grissmer & Kirby, 1997). Secondly, unemployment rates moderate the relation between job satisfaction and employee turnover (Carsten & Spector, 1987); when there are fewer options to change jobs, turnover rates tend to shrink. This affects studies relying on intentions to leave as a relevant factor for turnover. Finally, teacher turnover rates may vary between cohorts, as a consequence of policy shocks (e.g., budget cuts) which again are forces different from the school and from the teachers (Singer & Willett, 1988). Thus, the present research offers just one small step forward in understanding the contribution of individual and contextual factors in predicting teacher turnover, and our results must in the future be situated within a more comprehensive model of explanatory factors.

Figures and Tables Paper 1

Figure 1.1 The Job Demands-Resource Model (Bakker & Demerouti, 2007)

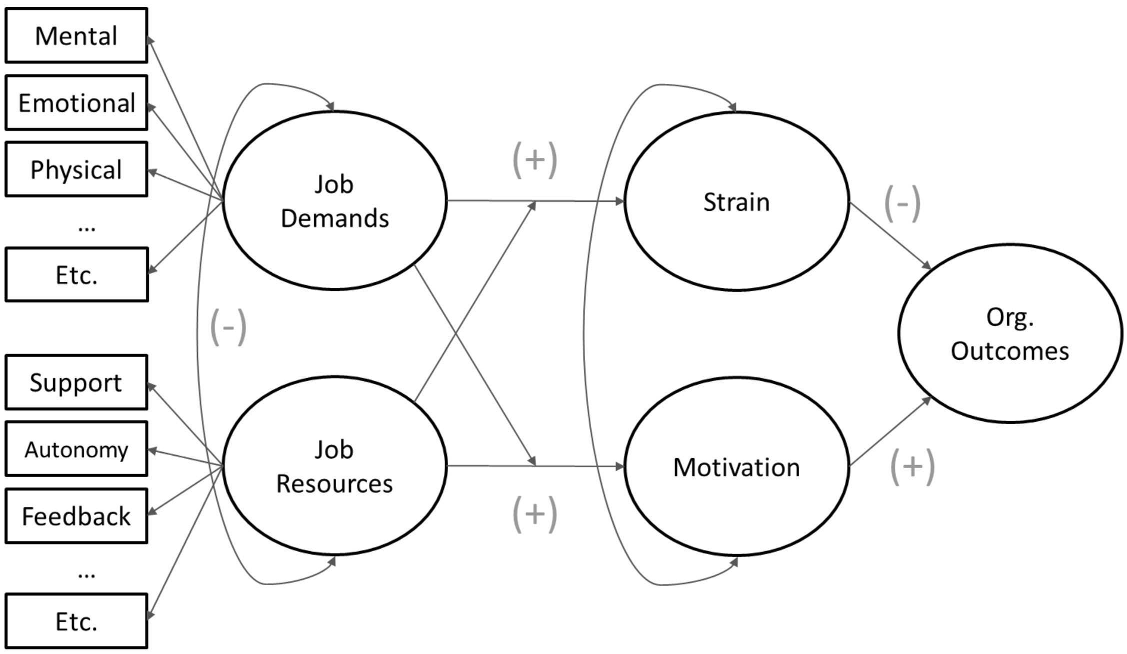


Table 1.1 Sample descriptives

	Model 1: All Schools			Model 2: Public Schools		
	Mean	*SE*	N	*Mean*	*SE*	N
School Differences						
Rural	.10	.00	7458	.11	.00	6570
School Size[c]	32.75	.32	7458	29.69	.28	6570
Class Size						
< 25 students per class	.35	.01	7427	.32	.01	6548
25 to 33 students per class	.28	.01	7427	.28	.01	6548
33 or more	.36	.01	7427	.41	.01	6548
School administration						
Public Schools	.22	.00	7458	.25	.01	6570
Subsidized Private Schools	.66	.01	7458	.75	.01	6570
Private Schools	.12	.00	7458	.00	.00	6570
School SES						
0% to 50% FSM students	.14	.00	7458	.16	.00	6570
51% to 70% FSM students	.23	.00	7458	.26	.01	6570
71% to 80% FSM students	.15	.00	7458	.17	.00	6570
81% to 100% FSM students	.23	.00	7458	.27	.01	6570
Life course factors						
Age at Entry						
Less than 30 years	.74	.01	7458	.74	.01	6570
30-40 years	.17	.00	7458	.18	.00	6570
More than 40 Years	.09	.00	7458	.09	.00	6570
Sex						
Female	.76	.00	7458	.75	.01	6570
Male	.24	.00	7458	.25	.01	6570
Working conditions						
Indefinite Contract	.28	.01	7458	.25	.01	6570
Other Contracts	.72	.01	7458	.75	.01	6570
Contracted Hours						
<29 hours a week	.44	.01	7458	.44	.01	6570
30 to 35 hours a week	.28	.01	7458	.28	.01	6570
36 hours or more	.28	.01	7458	.28	.01	6570
Number of Schools[c]	1.06	.00	7458	1.06	.00	6570
Teachers Role						
Primary School Teachers	.55	.01	7458	.57	.01	6570
Nursery Educators	.14	.00	7458	.12	.00	6570
Special Needs teachers	.09	.00	7458	.10	.00	6570
Secondary teachers	.22	.00	7458	.22	.01	6570
Sample Estimated Hazard						
Hazard t1	.34	.01	7458	.35	.01	6570
Hazard t2	.24	.01	4886	.25	.01	4266
Hazard t3	.17	.01	3693	.17	.01	3212
Hazard t4	.20	.01	3050	.18	.01	2660
Hazard t5	.13	.01	2444	.12	.01	2183
Hazard t6	.14	.01	2119	.14	.01	1916

Notes: Means and Standard Errors (SE) of all variables. All included variables are dummy coded variables, their expresses percentages; with the exemption of School Size and Number of schools, which are continuous (c).

Table 1.2 Fit Statistics of survival Estimates

	All Schools		Public Schools	
	Model 1		Model 2	
	Null model	*Covariates*	*Null model*	*Covariates*
Log likelihood	-12500.7	-12350.2	-10901.6	-10746.6
Scaling Correction Factor	1.2105	1.2259	1.151	1.2318
BIC	25054.91	24896.45	21855.88	21704.17
Free Parameters	6	22	6	24
Likelihood Ratio Test		244.431		246.18
df		16		18
p		0		0

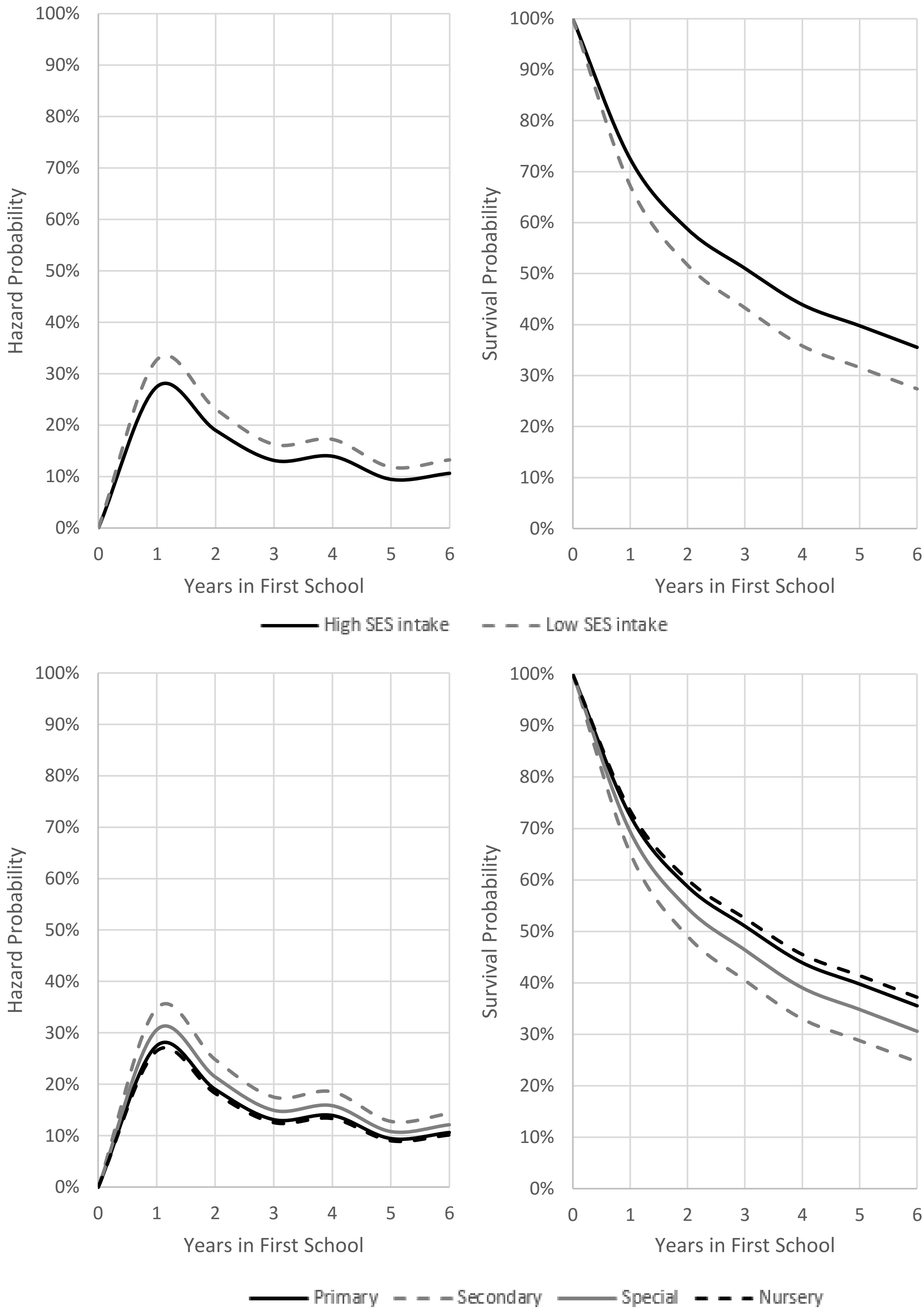
Table 1.3 Survival Estimates of duration of first teaching job

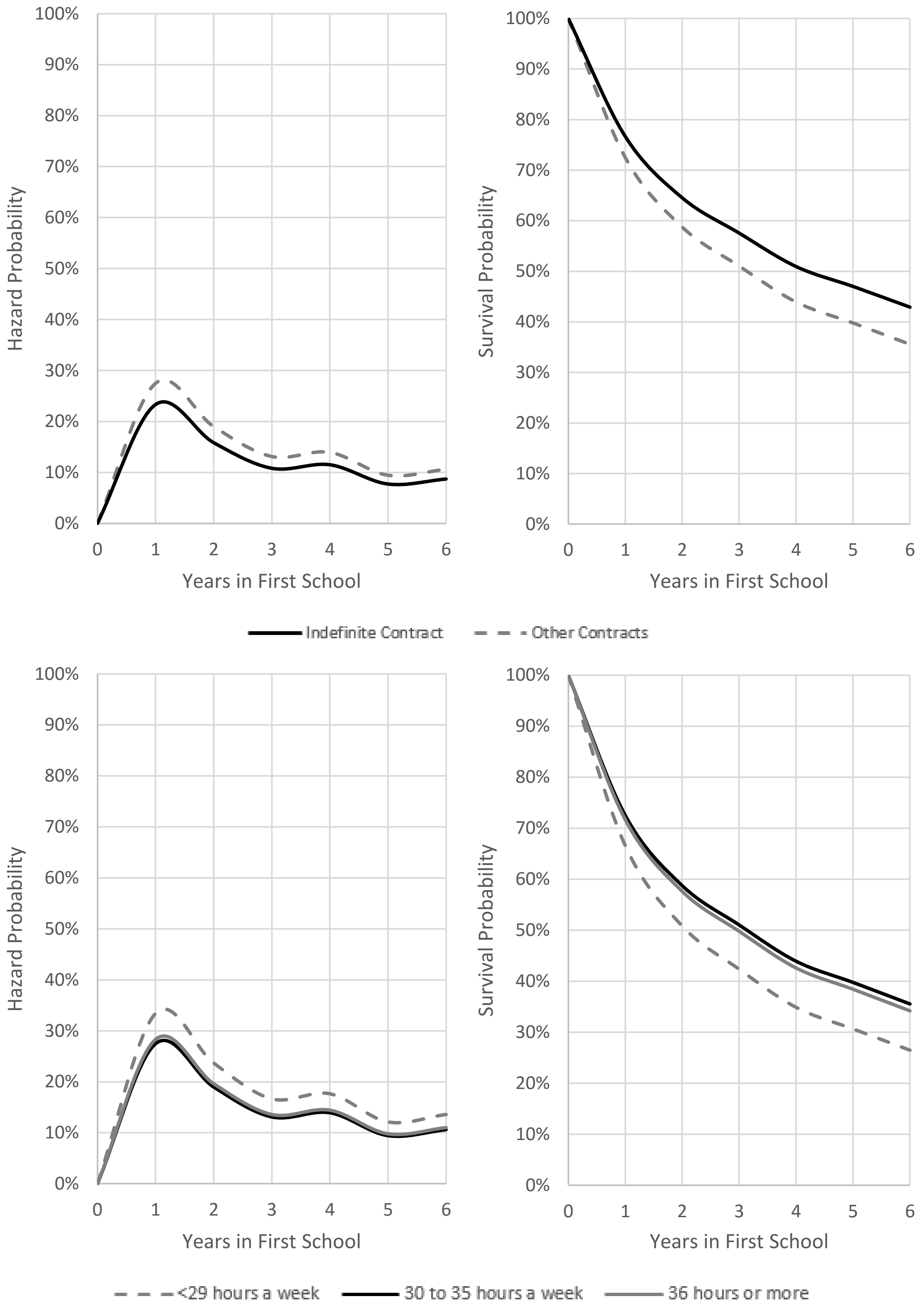
	All Schools				Public Schools			
	Model 1				Model 2			
	b		*SE*	*hOR*	*b*		*SE*	*hOR*
School Differences								
Rural	.13	*	(.06)	1.14	.04		(.07)	1.04
School Size	.00		(.00)	1.00	.00		(.00)	1.00
Class Size								
< 25 students per class	.13	*	(.05)	1.14	.15	*	(.06)	1.16
25 to 33 students per class	-----	----	-----	-----	-----	----	-----	-----
33 or more	-.13	*	(.05)	.88	-.12	*	(.05)	.89
School administration								
Public Schools	-----	----	-----	-----	-----	----	-----	-----
Subsidized Private Schools	.02		(.05)	1.02	.11	*	(.05)	1.12
Private Schools	.15	*	(.08)	1.16				
School SES								
0% to 50% IVE students					-----	----	-----	-----
51% to 70% IVE students					-.01		(.06)	.99
71% to 80% IVE students					.05		(.07)	1.05
81% to 100% IVE students					.25	**	(.07)	1.28
Life course factors								
Age at Entry								
Less than 30 years	.12	**	(.04)	1.13	.14	**	(.05)	1.15
30-40 years	-----	----	-----	-----	-----	----	-----	-----
More than 40 Years	.06		(.07)	1.06	.06		(.07)	1.06
Sex								
Female	-.11	**	(.04)	.90	-.12	**	(.04)	.89
Male	-----	----	-----	-----	-----	----	-----	-----
Working conditions								
Indefinite Contract	-.22	**	(.04)	.80	-.22	**	(.04)	.80
Other Contracts	-----	----	-----	-----	-----	----	-----	-----
Contracted Hours								
<29 hours a week	.27	**	(.04)	1.31	.28	**	(.05)	1.32
30 to 35 hours a week	-----	----	-----	-----	-----	----	-----	-----
36 hours or more	.07		(.05)	1.07	.06		(.05)	1.06
Number of Schools	.04		(.06)	1.04	.04		(.06)	1.04
Teachers Role								
Primary School Teachers	-----	----	-----	-----	-----	----	-----	-----
Nursery Educators	-.04		(.05)	.96	-.05		(.06)	.95
Special Needs teachers	.09		(.08)	1.09	.15		(.10)	1.16
Secondary teachers	.32	**	(.04)	1.38	.34	**	(.05)	1.40
Thresholds	*b*		*SE*	*h*	*b*		*SE*	*h*
Hazard t1	.85	**	(.08)	.30	.97	**	(.09)	.27
Hazard t2	1.32	**	(.08)	.21	1.45	**	(.10)	.19
Hazard t3	1.74	**	(.09)	.15	1.89	**	(.10)	.13
Hazard t4	1.55	**	(.09)	.18	1.82	**	(.10)	.14
Hazard t5	2.05	**	(.10)	.11	2.26	**	(.11)	.09
Hazard t6	1.96	**	(.09)	.12	2.13	**	(.11)	.11

* $p < .05$, ** $p < .01$

Notes: *b* = unstandardized estimates of maximum likelihood with robust errors, *hOR* = hazard odds ratio of teacher turnover; in the case of Thresholds these are hazard rates in probability scale, '-----'indicates reference category for categorical variables.

Figure 1.2 Fitted survivor and hazard probability describing the risk of leaving the first school contract, by SES intake and Teacher Role, Contract Status and Hours per week





Note: Fitted survivor and hazard function from Model 2, publicly funded schools, plotting the main effect of SES where High SES is 0% to 50% IVE students, and Low SES is 80% to 100% IVE student's intake. The black line in each graph is the reference category for each comparison.

Table 1.4 Study 2 Sample Descriptive Estimates

	Mean	(SE)	VIF	N	hOR		95% CI LL	95% CI UL
Teacher Turnover	.14	.00		10895				
School Differences								
SES: no IVE% students [d*]	.00	.00		10839				
SES: IVE 1% to 50% students [d*]	.05	.01		10839	.75		.54	1.06
SES: IVE 51% to 70% students [d]	.26	.01	4.73	10839	.94		.79	1.11
SES: IVE 71% to 80% students [d]	.31	.01	5.60	10839	.99		.85	1.15
SES: IVE 81% to 100% students [d]	.38	.01	6.16	10839	1.12		.97	1.29
Rural [d]	.28	.01	1.67	10655	1.00		.86	1.17
School Size (N. of Teachers) [c]	28.45	.50	1.79	10655	1.00	*	.99	1.00
Aver. <25 students per class [d]	.35	.01	1.89	10894	1.13		.98	1.30
Aver. 25 to 33 students per class [d*]	.35	.01		10894	.93		.80	1.08
Aver. 33 or more students per class [d]	.30	.01	1.45	10894	.95		.81	1.11
Acad. School Results (Math) [c]	.08	.02	1.26	9903	.87	**	.80	.95
Life course factors								
Age: Young (< 30 years) [d]	.16	.00	2.15	10884	2.27	**	1.99	2.59
Age: Mature (30-40 years) [d]	.18	.00	.00	10884	1.61	**	1.42	1.83
Age: Older (40 years or more) [d]	.65	.01	2.67	10884	.42	**	.37	.47
Exp: Novice (0 to 3 years) [d]	.20	.01	3.71	10611	2.39	**	2.10	2.72
Exp: Experienced (4 to 9 years) [d*]	.14	.00	2.27	10611	1.74	**	1.51	2.01
Exp: Senior (10 to 20 years) [d]	.21	.00	1.52	10611	.94		.82	1.08
Exp: Veteran (20 or more years) [d]	.45	.01	.00	10611	.37	**	.32	.42
Female [d]	.69	.01	1.09	10895	.80	**	.71	.90

* p <.05, ** p < .01

Notes: Mean = mean descriptive estimates for continuous and binary covariates, SE = clustered standard errors, VIF=variance inflation factor, hOR = hazard odds ratio of teacher turnover for a model including only the corresponding covariate with clustered errors, Exp = experience, Role = teaching role, Qual = qualifications, LL = lower limit, UL = upper limit, N = indicates the amount of valid cases per variable

[d] dummy coded variables, reference category use in models is marked with *

[c] non-binary variables (centred to the grand mean to estimate single covariate estimates)

Table 1.5 Study 2 Sample Descriptive Estimates (… **continuation**)

	Mean	(SE)	VIF	N	hOR	95% CI LL	95% CI UL
Working conditions							
Contract (indefinite)	.66	.01	1.73	10655	.33 **	.29	.38
Contracted Hours [c]	35.01	.10	1.31	10655	.95 **	.94	.96
Number of Schools [c]	1.12	.01	1.22	10655	1.57 **	1.37	1.81
Workload [c]	4.00	.01	1.06	10403	.95 *	.91	.99
Teachers Qualifications							
Role: Nursery [d]	.03	.00	1.05	10895	1.49 *	1.10	2.02
Role: Elementary [d*]	.30	.01		10895	.65 **	.57	.74
Role: Middle School [d*]	.47	.01		10895	.85 **	.75	.95
Role: High School [d]	.15	.01	1.41	10895	1.73 **	1.45	2.06
Role: Special Needs [d]	.06	.00	1.08	10895	1.77 **	1.43	2.19
NTES (high performance) [d]	.64	.01	1.09	10895	.50 **	.45	.57
Qual.: PhD [d]	.00	.00	1.01	10611	.65	.15	2.82
Qual.: Master [d]	.06	.00	1.04	10611	1.45 **	1.18	1.78
Qual.: Post Title [d]	.46	.01	1.13	10611	.64 **	.57	.72
Qual.: Diploma [d]	.10	.00	1.05	10611	.88	.73	1.06
Qual.: Training > 60 hh [d]	.44	.01	1.45	10611	.53 **	.47	.60
Qual.: Training > 60 hh (subject) [d]	.12	.00	1.22	10611	.78 **	.64	.94
Prior Intentions							
Intentions to Move [c]	2.31	.02	1.05	7611	1.2 **	1.15	1.24
Organizational Factors							
School leadership [c]	4.05	.01	1.99	10462	0.85 **	.80	.90
Positive School relations [c]	4.19	.01	1.60	10498	0.86 **	.81	.92
Academic monitoring [c]	4.22	.01	1.75	10349	0.96	.90	1.03
Staff relations is a priority [c]	4.20	.01	2.22	10309	0.86 **	.81	.92

* p <.05, ** p < .01

Notes: Mean = mean descriptive estimates for continuous and binary covariates, SE = clustered standard errors, VIF=variance inflation factor, hOR = hazard odds ratio of teacher turnover for a model including only the corresponding covariate with clustered errors, Exp = experience, Role = teaching role, Qual = qualifications, LL = lower limit, UL = upper limit, N = indicates the amount of valid cases per variable
[d] dummy coded variables, reference category use in models is marked with *
[c] non-binary variables (centred to the grand mean to estimate single covariate estimates)

Table 1.6 Model fit

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
-2 Loglikehood	7715.78	7092.20	6812.01	6664.41	4751.68	4644.01
Scaling Correction Factor	1.75	1.47	1.43	1.31	1.19	1.17
Parameters	9.00	15.00	19.00	29.00	30.00	34.00
BIC	7798.38	7229.45	6985.56	6929.30	5015.90	4943.06
Pseudo R^2	.01	.11	.14	.17	.20	.21
Δ Pseudo R^2	.00	.10	.03	.03	.03	.01
N clusters	2576	2526	2517	2517	2303	2293
N observations	9677	9412	9266	9266	6684	6605

Table 1.7 Logistic regression analysis of the likelihood of teacher turnover

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*hOR*
Intercept	1.99 **	-.22	1.45 **	-.24	1.35 **	-.25	1.09 **	-.26	1.19 **	-.28	1.24 **	-.28	.22
School Differences													
SES: IVE 1% to 50% students	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
SES: IVE 51% to 70% students	.18	-.22	.28	-.22	.22	-.23	.21	-.24	.28	-.26	.31	-.26	1.36
SES: IVE 71% to 80% students	.18	-.22	.30	-.22	.23	-.24	.20	-.24	.31	-.26	.32	-.26	1.38
SES: IVE 81% to 100% students	.20	-.23	.31	-.23	.24	-.25	.20	-.25	.24	-.26	.25	-.27	1.28
Urban	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
Rural	-.25 *	-.11	-.24 *	-.12	-.16	-.12	-.17	-.12	-.02	-.13	.01	-.13	1.01
School Size (N. of Teachers)	.00	.00	.00	.00	-.01 *	.00	-.01 **	.00	-.01	.00	-.01 *	.00	.99
Aver. <25 students per class	.03	-.11	.02	-.11	.01	-.12	.04	-.12	-.06	-.13	-.09	-.13	.91
Aver. 25 to 33 students per class	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
Aver. 33 or more students per class	.07	-.10	.09	-.10	.10	-.10	.03	-.10	.04	-.11	.03	-.11	1.03
Acad. School Results (Math)	-.12 *	-.05	-.10	-.05	-.10	-.06	-.06	-.06	-.04	-.05	-.02	-.05	.98
Life course factors													
Age: Young (< 30 years)			.10	-.10	.08	-.10	.03	-.10	-.04	-.12	-.03	-.12	.97
Age: Mature (30-40 years)			--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
Age: Older (40 years or more)			-.16	-.09	-.16	-.10	-.11	-.10	-.10	-.12	-.09	-.12	.91
Exp: Novice (0 to 3 years)			.15	-.09	.06	-.10	.05	-.10	.19	-.12	.23	-.12	1.25
Exp: Experienced (4 to 9 years)			--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
Exp: Senior (10 to 20 years)			-.43 **	-.10	-.21	-.11	-.22 *	-.11	-.19	-.13	-.19	-.13	.83
Exp: Veteran (20 or more years)			-1.08 **	-.12	-.62 **	-.13	-.58 **	-.13	-.54 **	-.16	-.54 **	-.16	.58
Female			-.21 **	-.07	-.16 *	-.07	-.08	-.08	-.14	-.09	-.11	-.09	.89
Male			--- ---	---	--- ---	---	--- ---	---	--- ---	---	--- ---	---	---

Notes: *b* = unstandardized estimates of maximum likelihood with robust errors, *hOR* = hazard odds ratio of teacher turnover, Exp = experience, Role = teaching role, Qual = qualifications

* *p* <.05, ** *p* < .01

Table 1.8 Logistic regression analysis of the likelihood of teacher turnover (continuation)

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*hOR*
Working conditions													
Indefinite Contract					-.59 **	-.09	-.51 **	-.09	-.59 **	-.11	-.60 **	-.11	.55
Other Contract					--- ---	---	--- ---	---	--- ---	---	--- ---	---	---
Contracted Hours					-.03 **	-.01	-.03 **	-.01	-.03 **	-.01	-.03 **	-.01	.97
Number of Schools					.10	-.08	.06	-.08	.16	-.10	.19	-.10	1.21
Workload					.02	-.03	.04	-.03	.02	-.03	.00	-.03	1.00
Teachers Qualifications													
NTES (low performance)							--- ---	---	--- ---	---	--- ---	---	---
NTES (high performance)							-.52 **	-.07	-.54 **	-.08	-.52 **	-.08	.60
Primary School Teachers							--- ---	---	--- ---	---	--- ---	---	---
Nursery teachers							.18	-.19	-.05	-.23	-.03	-.23	.97
Special Needs teachers							.72 **	-.13	.71 **	-.15	.68 **	-.15	1.98
Secondary School teachers							.32 **	-.11	.22	-.12	.17	-.12	1.18
PhD							-1.45	-1.09	-1.07	-1.16	-1.01	-1.11	.36
Master							.50 **	-.12	.45 **	-.14	.42 **	-.14	1.52
Post Title							-.18 *	-.07	-.21 *	-.08	-.21 *	-.09	.81
Diploma							.00	-.11	.10	-.13	.10	-.13	1.11
Training > 60 hh							-.22 **	-.08	-.21 *	-.10	-.23 *	-.11	.80
Training > 60 hh (subject)							.01	-.12	.11	-.14	.07	-.15	1.07
No additional qualification							--- ---	---	--- ---	---	--- ---	---	---
Withdrawal Intention													
Intentions to Move									.16 **	-.02	.14 **	-.02	1.14
Organizational Factors													
School leadership											-.15 **	-.06	.86
Positive School relations											-.08	-.05	.92
Academic monitoring											.12	-.07	1.13
Staff relations is a priority											-.15 *	-.06	.86

Notes: *b* = unstandardized estimates of maximum likelihood with robust errors, *hOR* = hazard odds ratio of teacher turnover, Exp = experience, Role = teaching role, Qual = qualifications. * p <.05, ** p < .01

Table 1.9 Logistic regression estimates, by groups of Novice and Non-Novice teachers

	Novice (0-3 years)				Non-Novice (4> years)			
	b		*SE*	*hOR*	*b*		*SE*	*hOR*
Intercept	1.42	**	-.37	.20	1.03	**	-.37	.26
School Differences								
SES: IVE 1% to 50% students	---	---	---	---	---	---	---	---
SES: IVE 51% to 70% students	.31		-.34	1.36	.29		-.35	1.33
SES: IVE 71% to 80% students	.57		-.34	1.76	.18		-.36	1.19
SES: IVE 81% to 100% students	.57		-.35	1.76	.04		-.37	1.04
Urban	---	---	---	---	---	---	---	---
Rural	.13		-.22	1.14	-.10		-.17	.91
School Size (N. of Teachers)	-.01		-.01	.99	-.01	*	-.01	.99
Aver. <25 students per class	-.12		-.21	.89	-.02		-.17	.98
Aver. 25 to 33 students per class	---	---	---	---	---	---	---	---
Aver. 33 or more students per class	.18		-.17	1.19	-.04		-.14	.96
Acad. School Results (Math)	.04		-.08	1.04	-.05		-.07	.96
Life course factors								
Age: Young (< 30 years)	-.07		-.16	.93	.06		-.19	1.06
Age: Mature (30-40 years)	---	---	---	---	---	---	---	---
Age: Older (40 years or more)	-.24		-.25	.79	-.25	*	-.12	.78
Female	-.01		-.15	.99	-.16		-.11	.86
Male	---	---	---	---	---	---	---	---
Working conditions								
Indefinite Contract	-.08		-.18	.92	-.93	**	-.12	.39
Other Contract	---	---	---	---	---	---	---	---
Contracted Hours	-.03	**	-.01	.98	-.03	**	-.01	.97
Number of Schools	.01		-.14	1.01	.31	*	-.13	1.36
Workload	.08		-.05	1.09	-.04		-.04	.96
Teachers Qualifications								
NTES (low performance)	---	---	---	---	---	---	---	---
NTES (high performance)	-.39	**	-.13	.68	-.58	**	-.11	.56
Primary School Teachers	---	---	---	---	---	---	---	---
Nursery teachers	.30		-.35	1.34	-.32		-.34	.73
Special Needs teachers	.80	**	-.28	2.23	.59	**	-.19	1.81
Secondary School teachers	.19		-.17	1.21	.17		-.16	1.18
PhD	.98		-1.32	2.67	-8.53	**	-.55	.00
Master	-.03		-.34	.97	.55	**	-.16	1.73
Post Title	-.20		-.16	.82	-.23	*	-.10	.79
Diploma	.49		-.30	1.64	.01		-.15	1.01
Training > 60 hh	-.37		-.25	.69	-.24	*	-.12	.79
Training > 60 hh (subject)	.39		-.35	1.47	.00		-.16	1.00
No additional qualification	---	---	---	---	---	---	---	---
Withdrawal Intention								
Intentions to Move	.07		-.04	1.07	.17	**	-.03	1.19
Organizational Factors								
School leadership	-.24	*	-.10	.79	-.13		-.07	.88
Positive School relations	.04		-.09	1.04	-.14	*	-.07	.87
Academic monitoring	.26	*	-.12	1.29	.05		-.08	1.05
Staff relations is a priority	-.21	*	-.10	.81	-.10		-.07	.91

Notes: *b* = unstandardized estimates of maximum likelihood with robust errors, *hOR* = hazard odds ratio of teacher turnover, Exp = experience, Role = teaching role, Qual = qualifications. * p <.05, ** p < .01

Table 1.10 Model fit for Novice and Non-Novice Teachers

	Novice	Non-Novice
-2 Loglikehood	1541.54	3054.03
Scaling Correction Factor	1.07	1.14
Parameters	31.00	31.00
BIC	1767.42	3318.95
Pseudo R^2	.10	.23
Δ Pseudo R^2		
N clusters	1014	2070
N observations	1460	5145

Note: model fit indexes were retrieved for both models estimated separately.

Paper 2. Teacher turnover intentions: Contextual effects and school climate predictors among secondary school teachers in Chile

Diego Carrasco
Robin Banerjee

University of Sussex, UK

Reference:

Carrasco, D., & Banerjee, R (2015). Teacher turnover intentions: Contextual effects and school climate predictors among secondary school teachers in Chile. *Unpublished manuscript.*

Abstract

A multilevel structural equation model was fitted to estimate the indirect relations of school intake characteristics and school climate factors with turnover intentions (intentions to quit the school, and intentions to leave the profession), via work satisfaction. The model was fitted among the Chilean population of lower secondary school teachers, from the TALIS (2013) teacher international survey. Positive teacher-student relations and classroom discipline were found to be protective factors against turnover intentions, with indirect effects via job satisfaction. Additionally, school climate factors had contextual effects: teachers in schools with greater classroom discipline presented teachers with lower intentions to quit; even after controlling for individual variations in job satisfaction and personal experience of classroom discipline. Results are discussed in relation to policies for teacher retention.

Keywords: teacher turnover, turnover intentions, teacher retention, job satisfaction.

Teacher turnover intentions, contextual effects and school climate predictors among secondary school teachers in Chile

Unequal educational systems, such as the system found in Chile (OECD, 2012), often suffer from unequal allocation of qualified teachers (Akiba et al., 2007), which partially explains inequities of academic results among students (Liu et al., 2015). This unequal distribution of teachers between schools is necessarily a function of how teachers take their first jobs, as well as how much time teachers remain in these schools and/or turnover from these jobs (Lankford, Loeb, & Wyckoff, 2002; Podgursky, Monroe, & Watson, 2004).

However, much of the policy efforts to change the distribution of teachers have a focus on recruitment, with much less attention to factors that promote retention. In other words, there is an unbalanced consideration of how to get teachers into certain jobs versus how to keep teachers in those jobs (Ingersoll, 2002). This is also the case in Chile, in which the focus tends to centre on “how to attract” more talented/qualified teachers to more vulnerable schools (e.g., Eyzaguirre, 2015). Examples of initiatives in this line are “Enseña Chile” and “Beca Vocación de Profesor”. *Enseña Chile* is a similar program to Teach for America in the US, and Teach First in UK, whereby people with experience of other professions are allowed to teach in schools via a shorter training path (Alfonso, Santiago, & Bassi, 2010). *Beca Vocación de Profesor* encourages high academic ability students to get a teacher degree with a merit based scholarship (MINEDUC, 2014). These contrast with a noticeable lack of policies to strength teacher induction and professional development of new teachers in order to buffer against teacher turnover (Avalos & Aylwin, 2007; Avalos & Sevilla, 2013; Avalos, 2009).

The uneven distribution starts from the beginning of teachers’ careers (Cabezas et al., 2011; Ortúzar, Flores, Milesi, & Cox, 2009; Ruffinelli & Guerrero, 2009). In Chile, recently graduated teachers distribute asymmetrically between schools. Teachers’ university entry exams scores are correlated to their schools’ socioeconomic status. Teachers with higher scores tend to be hired by private schools, and teachers with lower scores tend to be hired by public schools (Avalos, 2009). Similarly, teachers with low scores on their exit exam INICIA (Rivero, Hurtado, & San Martín, 2015) and first generation scholars, tend be hired on a higher proportion in public schools and schools with lower socioeconomic status (SES) intakes (Meckes & Bascopé, 2012). Furthermore, teachers’ uneven distribution is hindered by ensuing variations in turnover rates. Less qualified teachers are more likely to stay in low-performing low-income schools, in comparison to their more qualified counterparts. The contrary case is observed in schools with higher academic performance and higher SES students intake: less qualified teachers are less likely to stay in those schools, in contrast to their peers (Rivero, 2015). Hence, teachers' career trajectories appear to deepen uneven distributions.

Yet teacher retention is as important as recruitment to explain teacher distribution between schools. Focusing only on recruitment may leave this problem unsolved, because attracting more teachers does not address the retention side of the problem (Ingersoll & Smith, 2003). Attempting to change the distribution of teachers only by attraction may apparently work, but if teachers are not able to stay for long, this is not a solution on its own.

By studying teachers' distributions between schools and teachers' careers paths, we know more about if they leave, when they leave, and which career trajectories teachers often take (Cabezas et al., 2011; Rivero, 2015; Willett & Singer, 1991). However, we know less about why they leave. If we want to address the other side of the teacher-sorting problem, we need to address teachers' likelihood to quit, before it happens. The study of teacher turnover intentions may shed light in this direction.

Teacher turnover in Chile

Few studies have addressed the problem of teacher turnover within the Chilean population. Valenzuela and Sevilla (2013) estimated the rates of attrition of new teachers, between 2000-2008. They found teachers present similar rates of attrition to the US and UK, after the first teaching spell (9%-12% for years 2000-2004). These rates seem to be higher for years 2005 to 2008, reaching 18-20% of attrition for new teachers. According to the authors, this positioned Chile with the highest teacher attrition among OECD countries. Cabeza and colleagues (Cabezas et al., 2011) studied the relation between teachers' characteristics on the one hand and rates of teacher turnover and teacher attrition on the other. These authors found that being male, from low SES background, with less experience, from nursery or secondary school roles, and with higher credentials (Masters and/or PhD degrees) predicted higher levels of turnover. Rivero (2013) addresses the question of what teacher working conditions were related to teachers' mobility between schools (turnover), and leaving the public school system altogether (attrition), between 2009 and 2010. The author found that teacher influence, staff relations, facilities, principal support and school safety were all predictors of teacher turnover from the public system (attrition), and school safety and principal support were the most relevant factors that drive teachers' choices. Zamora (2009) studied the relationship between teacher commitment and intentions to leave, using Allen and Meyer's organizational commitment framework (N. J. Allen & Meyer, 1990; Meyer & Allen, 1991). The author found that three components of organizational commitment (affective, continuance, and normative) predicted teacher turnover intentions.

Limitations of previous research

Similarly to the general literature on teacher turnover, the previous research conducted in Chile agrees on the relevance of school climate factors, referring in general to terms such as work climate (Valenzuela, 2013), labour climate (Cabezas et al., 2011), and general working conditions (teacher influence, principal support, staff relations, school safety, and academic climate) (Rivero, 2013). However, only a few studies have assessed the relative impact of these factors on teacher turnover and teacher turnover intentions, with the exception of Rivero (2013) and Zamora (2009), respectively. The present study aims to compare the relative effect of different school climate factors on teacher turnover intentions for the population of teachers.

Previous research does not contain an explanation of why school environments and working conditions are related to turnover. It provides a reasonable expectation of the direction of effects, but not a conceptual model that explains why these effects occur, with the exception of Zamora (2009), who uses an organizational framework. In general, previous literature on teacher turnover fails to make connections with the general literature of withdrawal behaviours from the organizational behaviour literature (Berry et al., 2012; Griffeth et al., 2000; Hom et al., 1992; Podsakoff et al., 2007). In the present study, we offer a theoretically specified model that provides an explanation for the estimated parameters.

A third limitation present in the previous literature pertains to model specification. An over-reliance on univariate approaches, i.e. traditional regression analysis, may hide more complex effects. This is manifest in the previous studies via the specification of all factors as directly affecting the outcomes under study. This kind of model specification, however, cannot account for the interrelation between predictors, thus failing to account for relations that are more complex, such as indirect effects.

Additionally, previous research does not disentangle the plausible difference of within and between effects, thus providing conflated estimates (K. J. Preacher et al., 2011a) or inconsistent estimates. Given that teachers are nested in schools, some relations may be general to all teachers regardless of the school environment, while others might be relations pertaining to differences between school environments, regardless of a teacher's particular experiences. On the one hand, if only between-school effects are included, and the within-school level is ignored, parameters for the former will be overestimated. Likewise, if between-school variance of covariates is ignored, within-school estimates will be inconsistent (Antonakis et al., 2010, 2014). Moreover, manifest mean aggregation of within level covariates to create cluster level effects, can lead to an underestimation of contextual effects, due to sampling error (Lüdtke et al., 2008; Pokropek, 2015; Televantou et al., 2015). Traditional regression approaches do not partition the effect of covariates of nested observations, but this is possible within regressions and multilevel models variants (McNeish & Stapleton, 2015; Stapleton, 2013). We therefore use a multilevel latent covariate model (Lüdtke et al., 2008, 2011) to overcome jointly these limitations.

Theoretical Framework

Turnover intentions precede actual turnover behaviour, under the view of voluntary turnover. Podsakoff and colleagues (Podsakoff et al., 2007) present a general model in which strains diminish job satisfaction, which in turn is connected to organizational commitment, turnover intentions and actual turnover. In a conceptual way, the general model can be expressed by the following sequence: strains → less satisfaction → less commitment → more withdrawal intentions → more turnover.

Similarly, the job demands and resources model (Hakanen et al., 2006) posits job demands as risk factors for work stress, and job resources as protective factors for work engagement. Both dimensions are related to organizational outcomes, including strain related outcomes (e.g., ill health, absenteeism, turnover) and motivational process outcomes (e.g., in role performance, extra role performance) (Bakker & Demerouti, 2007). Thus, both models converge on the general idea that strains are indirectly related to the turnover process.

In the present study, *school intake characteristics* are included as a strain factor. Previous literature establishes that schools with staffing difficulties tend to have intakes with higher proportions of students from minority backgrounds, low academic achievers, and low SES (Djonko-Moore, 2015; Guin, 2004; Horng, 2009). We used two *school climate factors* to depict demands and resources: positive teacher-student relations and student misbehaviour. Both factors have known positive and negative relations, respectively, to teacher turnover (Allensworth et al., 2009; Bakker et al., 2007; Kelly, 2004; Kukla-Acevedo, 2009).

Teacher's job satisfaction is included as a mediator, for all previous factors, predicting turnover intentions. Finally, two specific variables were used to measure *turnover intentions*: intentions to quit the school and intentions to leave the profession. The latter are withdrawal cognitions, reflecting negative attitudes towards the teaching career. In essence, it represents the reverse attitude to that of career commitment (Blau & Lunz, 1998). In contrasts, intentions to quit the school reflect a direct desire to leave the current job for a different workplace, while remaining in the same career path (Goddard & O'Brien, 2003; Hom et al., 1992; Podsakoff et al., 2007; van Breukelen, van der Vlist, & Steensma, 2004).

Teacher turnover intentions refer to the desire to quit or move from a job, regardless of whether or not the turnover actually happens. General meta-analytic estimates of associations between turnover intentions and actual quitting behaviour among different occupations show an r of .45, when measurement, sampling and variations of base rate turnover are taken into account (Griffeth et al., 2000). Turnover intentions, therefore, are an indirect way to measure teachers' likelihood to quit. However, it is an imperfect measure. The relation of teacher turnover intentions to actual turnover is moderated by available job positions (Ladd, 2011), unemployment rates (Hom et al., 1992) and non-voluntarily layoffs (Singer & Willett, 1988), all of which are exogenous factors to voluntarily turnover.

Conceptually, turnover intentions are considered the last step in a sequence of withdrawal cognitions that lead to turnover (Tett & Meyer, 1993). In essence, turnover intentions are the most proximal mediator, between working conditions, job satisfaction, commitment, and actual turnover (Hom et al., 1992; Kammeyer-Mueller, Wanberg, Glomb, & Ahlburg, 2005; Mitchell & Lee, 2001; Podsakoff et al., 2007; van Breukelen et al., 2004). Thus, in spite of its imperfections, it is considered a good proxy for actual turnover (Bothma & Roodt, 2013; Sousa-Poza & Henneberger, 2004).

The present study

In the present study, we address the following questions: a) What is the relative contribution of different school and teacher characteristics and school climate factors to teacher turnover intentions? b) what factors account for these relationships? and c) What is the role of school climate factors in these relations?

In order to assess the relative relevance of school climate factors onto teacher turnover intentions, we include a series of control variables frequently used in the literature on teacher turnover (e.g., school characteristics, social demographics, teacher qualifications, working conditions) and teacher turnover intentions are regressed on these, along with school climate factors. We present population estimates for each of these factors (see Tables 2.1 and 2.2). In a second step, we give preliminary evidence of the mediator effect of job satisfaction on teacher turnover intentions (see Tables 2.2 and 2.3). Finally, we fit a multilevel SEM (see Figure 2.1), to evaluate the proposed model, and assess the role of school climate factors. This model permits us to unravel within and between effects, estimate indirect effects (see Table 2.5) and contextual effects (see Table 2.6).

Method

Participants and data sources

Data from the Chilean participating sample was used from the Teaching and Learning International Survey from 2013 (TALIS 2013 for short) (OECD, 2014e). This is a Large Scale Assessment (L. Rutkowski et al., 2010), in which national representative samples are collected, using a two stage probability sample design, in order to provide representative samples of the population of teachers. A total of 34 countries and economies participated in this study from different regions, including: Chile and Mexico from Latin America; Canada and US from North-America; England, France, Netherlands from Europe; Japan and Korea from Asia - for an exhaustive list, see OECD (2013b). This survey targets a random sample of 200 school principals and 20 representative teachers within each schools, from lower secondary levels (ISCED=2, 7th to 9th grades) from mainstream education, with the purpose of achieving a hypothetical random sample of 433 teachers with a margin of error of 9.4% (OECD, 2013b). All analyses were conducted with the subpopulation of Chile within TALIS

2013, which held a nominal sample of 178 schools, and 1676 teachers. Population estimates[7] yield a mean age of 41.33, CI95% [40.44; 42.24] for teachers, with a majority of female teachers (62.79%, CI95% [60.12%; 65.38%]), from almost half of publicly funded schools (52.35%, CI95% [45.75%; 58.95%]). All estimates were calculated while accounting for the complex sample design of the study, to yield generalizable population estimates.

Variables

Variables used in the present study are describe in more detail in Appendix (Paper 2), Measure and Items section. This appendix includes the actual format of each question and score constructions procedures.

School Characteristics

Intake. Teachers answered what proportion of students have the following characteristics at the target class: low academic achievers, students with special needs, students with behavioural problems and students from low SES background. Answers were recorded using a scale from 1 to 5, with ordinal categories of: None, 1% to 10%, 11% to 30%, 31% to 60% and More than 60%. These answers were modelled as categorical factors with confirmatory factor analysis and standardized factors scores were retrieved. This procedure creates an index with a grand mean of zero for the population, whereby higher numbers indicate a higher proportion of students with more demanding characteristics.

School Administration. The Chilean school system is divided between three types of administrations: private schools, subsidized schools, and public schools. We created three dummy coded variables to identify each of this school types.

School Location. School location was addressed as an ordinal measure of urbanity, where higher numbers indicate school locations with higher density populations.

School Size. We used the total number of teachers within each school as a proxy for school size.

[7] These are descriptive results accounting for complex sample design, therefore their generalization domain is the subpopulation of Chilean teachers (Heeringa, West, & Berglund, 2009). A resampling technique was employed, BRR with 100 resamples, and Fay adjustment of .5. to produce these population descriptive estimates.

Teachers' Characteristics

Age. Teachers' age was available in years.

Experience. Teachers indicated how many years they have been working as a teacher.

Gender. Teacher's sex was computed as a dummy variable, in which Male is the reference category (with a zero value), and Females are identified with value of 1.

Teacher qualifications

Teacher Certification. Teachers indicated whether they have completed a teacher training program. Their answers were dummy coded, leaving teachers without teacher certification as a category of reference (yes=1, no=0).

Educational degree. Teachers answer what was their last educational degree. These answers were dummy coded, and we left "Bachelor's Degree" as a University Degree for reference.

Working Conditions

*Contract type. T*eachers indicated whether they were hired under a permanent contract or fixed term contract. This variable was coded as a dummy variable, leaving full contract as 1, and the rest of categories as reference (full=1, else=0).

In how many schools the teachers work. Teachers indicated if they worked in more than one school. Their answers were included as dummy variable (yes=1, no=0).

School Climate factors

Distributed Leadership. School principals rated the level of distributed leadership within the school, by indicating their level of agreement to assertions such as: 'This school provides staff with opportunities to participate actively in school decisions'. The answers to three similar items were factor scored to create a single index (see OECD, 2013b, p. 156). For the present study, we used the scoring provided by the OECD (2013b).

Instructional Leadership. School principals rated different affirmations regarding their involvement in the teaching practice (e.g., 'I took action to ensure that teachers take responsibility for improving their teaching skills'). Three items of this sort were factor scored to create a single scale by the OECD (2013b).

Mutual Respect. School principals rated their level of agreement with different items referring to teacher interpersonal relations (e.g., 'There is a mutual respect for colleagues' ideas', and 'School staff have an open discussion about difficulties'). Four of these items were factor scored by the OECD (2013b) to create a single scale.

Shortage. School principals indicated whether their school is affected by a shortage of qualified teachers and/or well performing teachers, using Likert scale of four levels, from 'Not at all' to 'A lot'. We used these answers as proxy indicators of previous teacher shortage at the school level.

School Delinquency. School principals rated how often different events of vandalism occur within the school. Four events were factor scored to create a single scale, including theft, intimidation, physical injury, and verbal abuse. We relied on the factor scoring provided by the OECD (2013b).

Positive Teacher-Student Relations. Teachers described the quality of interpersonal relations between teachers and students at their school (e.g., 'In this school, teachers and students get on well with each other'). Their answers were linearly combined using their weighted robust likelihood estimated loadings and intercepts. These were scaled to have a mean of 10 and a standard deviation of 2 across countries by the OECD (OECD, 2014f). The higher the score, the more positive the interpersonal relations between students and teachers.

Classroom Discipline. Teachers responded to four affirmations expressing how disciplined their students were (e.g., 'I lose quite a lot of time because of students interrupting the lesson'). Similar to the previous revised scales, these answers were scaled to have a mean of 10 and a standard deviation of 2. The higher the score, the more discipline the students have in the target class.

Job Attitudes

Job Satisfaction. Teachers rated in general how satisfied they felt with their school environment. They answered, with a Likert type Scale of four levels, their level of agreement with the following affirmations: 'I enjoy working at this school', 'I would recommend my school as a good place to work' and 'All in all, I am satisfied with my job'. These answers were factor scored as categorical answers under a MLR estimator to create a single scale.

Intentions to leave the profession. Similar to the previous factor, teachers indicated how satisfied they felt with their profession, using a Likert type scale of four levels. They responded to these items: 'The advantages of being a teacher clearly outweigh the disadvantages' (reversed), 'If I could decide again, I would still choose to work as a teacher' (reversed) and 'I wonder whether it would have been better to choose another profession'. These answers were factor scored as categorical observations under a MLR estimator to create a single index.

Intentions to quit. Teachers indicated their level of agreement with the following item: 'I would like to change to another school if that were possible', with a Likert type scale of four levels. This measure was included in the models as a single item.

Analysis plan

Derived scales. We first of all estimated the measurement model of Work Environment Satisfaction and Profession Satisfaction from the item scales set out in TALIS 2013 (OECD, 2014f). Originally, these scales include the “intentions to quit the school” as facets of their measurement. However, these are more distal facets from these constructs, and according to the theory of planned behaviour models of turnover (van Breukelen et al., 2004) can be viewed as a different dimension altogether. The original measurement model, includes 8 items and two factors, and it does not fit the data well ($\chi^2(19)= 295.596$, $p<0.001$, BIC=26538.036). A different model estimation was used, in which all items were modelled as categorical indicators. This model fits the data better (BIC=23910.324). Finally, we fit an alternative measurement model, in which intentions to quit and career withdrawal cognitions are endogenous to job satisfaction. This final model fits the data slightly better (BIC=23910.321). Given our interest in understanding intentions to leave the workplace, we preferred the alternative model. The original and alternative models are displayed in Figure 2.2 in Appendix (Paper 2).

Likewise, we created a school intake scale (conceptualized to operate as a job strain), by deriving its factor scores from a confirmatory factor analysis (see Appendix (Paper 2), Intake). In this model, school intake characteristics reflect the proportion of low academic achievers, students with special needs, students with behavioural problems and students from low SES, as these are all taken as more demanding student characteristics (Bakker et al., 2007; Guin, 2004; Horng, 2009).

Estimation. The TALIS (2013) survey presents a complex sample design, with teachers nested in schools, with random observations and unequal probability of selection (survey weights). To get population estimates in Tables 2.1, 2.2 and 2.3, we accounted for the sampling design and derived standard errors via the Balanced

Repeated Replication (BRR) method (Jensen et al., 2012), using STATA (StataCorp, 2013) and its '*svr module*' (Winter, 2008). In Table 2.1, alongside descriptives of each covariate, we estimated all correlation coefficients to each of our outcomes of interest: job satisfaction, intentions to leave the profession and intentions to quit the school. Additionally, we estimated the variance inflation factor (VIF) of all included variables, to assess possible multicollinearity problems between variables. However, none of the variables were VIF<10, which is a common rule of thumb for concern (O'Brien, 2007).

In Table 2.2, we estimated regression coefficients to explain job satisfaction, and turnover intentions. In the last two columns of this table, we fit regression coefficients to predict turnover intentions, but this time included job satisfaction as a predictor. Next, in Table 2.3, we fit the same regression without control covariates, demonstrating that the key coefficients are largely consistent with or without the inclusion of more covariates in the model.

Figure 2.1 depicts the standardized estimates of a multilevel structural equation model with fixed slopes to estimate indirect effects (K. J. Preacher et al., 2011a, 2010), specifying schools as clusters. To account for sampling design within a multilevel framework, we partitioned survey weights in within and between components (Kim, Anderson, & Keller, 2013). We scaled these survey weights to effective sample size (Asparouhov, 2006; Rabe-Hesketh & Skrondal, 2006; Snijders & Bosker, 2012). This scaling weight method was preferred over other weight normalization methods, because it produced unbiased estimates of variances (Stapleton, 2002). We used MPLUS v7 (L. K. Muthén & Muthén, 2012) software for this purposes. Results in Tables 2.5 and 2.6 are post estimation results from the general model presented in Figure 2.1.

Model Sequence. We fit a structural equation model to operationalize the conceptual relationship of school intake → school climate factors → satisfaction → turnover intentions. This model resembles model 6 ($x_{ij} \Rightarrow m1_{ij} \Rightarrow m2_{ij} \Rightarrow y_{ij}$) from Hayes (2013, p. 446), with the difference that we are fitting this model within a multilevel latent covariate framework. Therefore, the same model is fitted for the within and between covariance matrixes (Hox, 2010; Ryu, 2015). In this model school intake is the exogenous factor, while the school climate factors (positive teacher student relations and classroom discipline) are the distal mediators, job satisfaction is the proximal mediator, and finally teacher turnover intentions (intentions to leave the profession and intentions to leave the school) are the outcomes. The fit of the model sequence is displayed in Table 2.4.

Results

Population Estimates

Demanding intake characteristics have a small negative relation to school satisfaction ($r = -.14$), and a positive relation to turnover intentions to leave the profession ($r = .17$), and intentions to quit the school ($r=.14$). When these estimates are jointly estimated including different covariates including school characteristics, teacher characteristics, working conditions and school climate factors, school intake characteristics cease to present relevant effects on Job Satisfaction and Intentions to quit. However, there is still an independent effect of this covariate on intentions to leave the profession ($B=.11$, $SE=.05$, $p<.05$), even when all control variables are taken into account.

Table 2.1: Descriptives, Population estimates

Some heterogeneity between teachers is observed: older teachers present more job satisfaction than younger teachers do ($B=.26$, $SE=.01$, $p<.01$), and more experienced teachers tend to be less satisfied with their job environments ($B=.19$, $SE=.01$, $p<.05$). The reverse is observed regarding intentions to leave the profession: older teachers are less likely to present withdrawal cognitions regarding the profession ($B=-.26$, $SE=.01$, $p<.01$), and teachers with more years of experience are more likely to have withdrawal thoughts regarding the profession ($B=-.25$, $SE=.01$, $p<.01$). The rest of the included covariates all present smaller effects ($B<.13$). In contrast, no significant effects are observed for teacher intentions to quit from these factors, besides the main predictive effects of positive teacher-student relations and classroom discipline described below.

TABLE 2: Regressions Estimates with control variables

Positive teacher-student relations and classroom discipline are both consistent predictors of all job attitudes in the expected directions: Teachers who experience positive interpersonal relations with their students have higher job satisfaction (B=.36, SE=.02, p<.001), and teachers who teach in more disciplined classrooms have higher job satisfaction (B=.20, SE=.02, p<.001). Similarly, teachers with more positive interpersonal relations with their students exhibit less intention to leave the profession (B=-.27, SE=.02, p<.001) and less intention to leave the current school (B=-.28, SE=.02, p<.001). Finally, teachers who teach in well-behaved classrooms have less intention to leave the profession (B=-.18, SE=.02, p<.001), and less intention to quit the school (B=-.21, SE=.02, p<.001). These estimates are largely consistent, whether we include or exclude the effects of the control variables (see Tables 2.2 and 2.3).

Limitations of the population estimates approach

Population estimates are informative regarding the general relation between a covariate, or a vector of covariates, and an outcome. These are estimated relations between factors as if teachers were randomly picked out of the population. However, as the model was specified, it is uninformative regarding differences between school environments. It corrects the fact that observations are nested within schools to avoid Type 1 error, yet it is silent with respect to the fact that teachers from the same school provide information of the same referent, thus their answers are also interpretable as school attributes or differences between school environments. This model specification, even though is correct for standard error estimation, is blind to the inter-relation between covariates.

Additionally, the inclusion of many covariates from the school level leads to a higher proportion of missing data. This restricts the number of observations with valid information to almost half of the sample. From 1676 cases, listwise deletion with controls yields 720 cases. As noted above, we compared the consistency of the main coefficients of interest, with and without control variables (see Tables 2.2 and 2.3). These estimates were largely consistent, and were all within the same confidence interval.

Table 2.3: Regressions Estimates without control variables

Thus, to overcome the limitations noted above, we estimated a multilevel structural equation model with fixed slopes (K. J. Preacher et al., 2011a), with manifest variables and aggregated latent means (Lüdtke et al., 2011).

Multilevel Structural Equation Model

The fitted model split the variance of all the variables in two parts: the within-school and between-school covariance matrix. In general, this is expressed in the formula: $\Sigma_T = \Sigma_w + \Sigma_B$ (Hox, 2010, p. 291), where the matrix covariance can be split as additive terms. The within part (Σ_w) consists of all relative differences from the latent mean per cluster for all variables. The between part (Σ_b) concentrates the total differences between latent means of all clusters. This is similar to multilevel models with centering within cluster (Enders & Tofighi, 2007), with the difference that measures in this model are centred to the latent cluster mean (Nagengast & Marsh, 2012; Ryu, 2015).

Figure 2.1: MSEM standardized estimates

Within estimates are pooled estimates of all observations for all teachers across schools. In contrast, between estimates are aggregated differences between schools. In this case, they express latent mean aggregated relations (Lüdtke et al., 2008; Ryu, 2015). Thus, within estimates are informative regarding the interrelations between factors for all teachers across schools, in spite of school differences. In other words, these parameters estimate the relationship across factors, for all relative differences between teachers within schools. Between estimates, on the other hand, are the expected effects between variables at the school level. These estimates speak of the general trend between factors as aggregated measures at the school level. Inferential tests were conducted onto the unstandardized estimates and onto the standardized estimates. These two, did not differ. To facilitate the interpretation, we report the standardized estimates.

For example, the nonsignificant $X_w \Rightarrow Y1_w$ coefficient (B=.04, SE=.03, p=.16) in Figure 2.1 indicates that the relative differences between teachers within school regarding their perceptions of the school's student characteristics are not strongly associated with the relative differences in their intentions to leave the profession. However, the $X_b \Rightarrow Y1_b$ coefficient (B=.30, SE=.13, p=.02) tells a different story. Schools that are broadly perceived to have a higher proportion of students with more demanding characteristics are shown to have teachers with higher levels of intention to leave the profession.

Assuming that the between part of the model is equal to the within part, so one can reduce the latter to the former, is known as the Ecological Fallacy. On the other hand, assuming that the within part can simply be reduced to the between part is known as the Atomistic Fallacy (Rabe-Hesketh & Skrondal, 2012). Multilevel and MSEM models, permits to estimates relations between factors at each level and avoid ecological and atomistic fallacies.

The fitted model in Figure 2.1 is a saturated model. Hence, it does not produce fit indexes. To assess if the inclusion of each covariate adds value to the fit of the model we use a model building approach (Kline, 2010). We estimated first a direct effect model, in which demanding intake characteristics are the only covariate with freely estimated effects (Model 0) on teacher job satisfaction, intentions to leave the profession and intentions to leave the school. We then include the specification of job satisfaction as a mediator (Model 1), followed by a model in which positive teacher-student relations are included as a distal mediator (Model 2), and finally a model in which classroom discipline is also specified as a distal mediator (Model 3). Each of these steps yields a smaller deviance and significant likelihood ratio test (see Table 2.4), thus supporting the inclusion of the specified relations at each step.

Table 2.4: Model Sequence

Explained variance. Intraclass correlation values for each variable under a null model were $Y1_{ij}$ = 8.6%, $Y2_{ij}$ = 10.9%, $M2_{ij}$ = 12.5%, $M1A_{ij}$ = 12.0%, $M1B_{ij}$ = 16.6%, and X_{ij} = 30.7%. Thus, most of the variance of the variables is concentrated at the within level, reflecting relative differences among teachers within schools, with less variance located as differences between schools. Our chosen model specification explains 26% the variance at the within part for intentions to quit the school ($Y2_w$), and 54% of teachers' intentions to leave the profession ($Y1_w$). The chosen factors are thus better at explaining within-school variations in withdrawal cognitions of leaving the profession than teacher intentions to quit the school. At the between level, this model is able to account for 76% of the variance in teachers' intentions to leave the school ($Y2_b$), and 60% of teachers' intentions to leave the profession ($Y1_b$). Overall, chosen factors explained a high proportion of the variance (R^2>50%) at the between level.

Between Estimates. There are two direct effects of school intake in the model, at the between level. Schools with a higher proportion of students with more demanding characteristics (low academic achievers, low SES, behavioural problems, and special needs), are also schools with teachers with higher mean levels of withdrawal cognitions to leave the profession (B=.30, SE=.13, p=.02). Similarly, schools with a higher proportion of students with these characteristics, show lower levels of classroom discipline as a school (B=.33, SE=.14, p=.02). Putting these paths together, we learn that school intake has an indirect effect on teacher intentions to leave the school, via school classroom discipline (B=.19, SE=.09, p=.03).

School climate factors, namely positive teacher-student relations and classroom discipline, are positive predictors of teachers' job satisfaction at the school level, accounting for R^2=57% of the between variance (see Figure 2.1). Positive teacher-student relations and teachers' withdrawal cognitions to leave the profession are indirectly related via job satisfaction ($M1A_b \Rightarrow M2_b \Rightarrow Y1_b$: B=-.21, SE=.09, p=.02, this indirect effect is about 40% of the total effect). Classroom discipline, however, has direct effects on teachers' intentions to quit the school (B=-.54, SE=.19, p<.01), whereby schools with better discipline have teachers with less intention to quit their job.

In general, schools with higher levels of teachers' job satisfaction, also have teachers with lower levels of intention to quit the school (B=-.33, SE=.17, p<.05), and less intention to leave the profession (B=-.43, SE=.27, p<.05).

Contextual Effects. Additionally, the model presents contextual effects. These are expected effects on the outcomes, which are independent of teacher characteristics or relative levels. Because MSEM partitions all estimates into within and between parts, and within estimates are latent mean centred, we can get contextual effects for any coefficient, by subtracting the between estimate from its within counterpart (e.g., Nagengast & Marsh, 2012). By this method, we found two contextual effects (see Table 2.6). Teachers in schools with higher than average classroom discipline, present less intention to quit than those in schools with average levels of classroom discipline, regardless of their own target class's level of classroom discipline (contextual $M1B_{ij} \Rightarrow Y2_{ij}$: B=-.48, E=-.17, SE=.08, p<.05). We also found that teachers in schools with higher than average job satisfaction are more likely to exhibit higher withdrawal cognitions than those in schools with average levels of school satisfaction, beyond the role played by their own individual job satisfaction levels (contextual $M2_{ij} \Rightarrow Y1_{ij}$: B=.26, E=.26, SE=.17, p<.05). This rather counterintuitive finding is a reverse contextual effect, in which the within estimate is larger than the between estimate, and thus the contextual effect goes into an opposite direction than that of the within and between estimates.

Within Estimates. Class intake characteristics have no significant effects on any of the variables (standardized estimates sizes of .06 to .02), except that teachers who perceive more demanding student intake characteristics tend to report less discipline in their classrooms (B=-.16, SE=.04, p<.001), relative to their colleagues in the same school.

Teachers' perceptions of classroom discipline are a positive predictor of their job satisfaction, relative to their colleagues (B=.16, SE=.04, p<.001). Likewise, positive teacher-student relations also contribute to variance in teachers' job satisfaction within schools (B=.32, SE=.03, p<.001). Additionally, classroom discipline presents a small direct effect on teachers' intentions to leave the profession (B=-.06, SE=.03, p<.05). Apart from this effect, all relationships between these factors and teacher turnover intentions are indirect via job satisfaction. Relative differences in teacher job satisfaction within schools are negative predictors of teacher turnover intentions, both the intentions to leave the profession ($M2_w \Rightarrow Y1_w$: B=-.71, SE=.02, p<.01), and the intentions to leave the school ($M2_w \Rightarrow Y2_w$: B=-.47, SE=.03, p<.01).

Indirect effects. School intake characteristics present small indirect effects only via classroom discipline to all endogenous variables. Class intake has an indirect effect on teacher job satisfaction ($X_w \Rightarrow M1C_w \Rightarrow M2_w$: B=-.03, SE=.01, p<.01, about 27% of the total effect), as well as indirect effects on teacher turnover intentions via the classroom discipline-job satisfaction link. In this way, a classroom composition with more demanding students has an indirect positive relationship with teacher withdrawal cognitions to leave the profession ($X_w \Rightarrow M1C_w \Rightarrow M2_w \Rightarrow Y1_w$: B=.02, SE=.01, p<.01, about 16% of the total effect) and to intentions to quit the school ($X_w \Rightarrow M1C_w \Rightarrow M2_w \Rightarrow Y2_w$: B=.01, SE=.00, p<.01, about 13% of the total effect).

Positive teacher-student relations and classroom discipline have significant indirect effects on teacher turnover intentions via job satisfaction. The former yields a negative indirect standardized effect of -.23 ($M1A_w \Rightarrow M2_w \Rightarrow Y1_w$: B=-.23, SE=.02, p<.01, accounting for 89% of the total effect) on teachers' intentions to leave the profession. Similarly, it also presents a negative indirect effect on teachers' intentions to quit the school, via job satisfaction ($M1A_w \Rightarrow M2_w \Rightarrow Y2_w$: B=-.15, SE=.02, p<.01, accounting for 77% of the total effect). Classroom discipline presents a small indirect effect on teachers' intentions to quit the school ($M1B_w \Rightarrow M2_w \Rightarrow Y2_w$: B=-.08, SE=.02, p<.01, accounting for 56% of the total effect), and presents small direct and indirect effects on teachers' intentions to leave the profession ($M1B_w \Rightarrow M2_w \Rightarrow Y1_w$: B=-.11, SE=.03, p<.01, accounting for 66% of the total effect).

Discussion

School climate factors are relevant explanations of teacher turnover intentions. Positive teacher-student relations and classroom discipline present standardized coefficients as large or larger as any other factor, compared to school differences and teacher characteristics. Furthermore, when job satisfaction is also included in the model as a predictor, it accounts for most of previous effects, whether we include or exclude control variables. Thus, these estimates are highly consistent. By inspecting the within and between estimates from Figure 2.1, we can unravel the role of school climate factors on teacher turnover intentions. First, school climate factors can moderate individual outcomes (see Table 2.6) independent of teachers' individual experiences. At a contextual level, school classroom discipline behaves as a protective factor for teacher turnover intentions to leave profession, regardless of teachers' characteristics; and school job satisfaction is, surprisingly, associated with larger withdrawal cognitions to leave the profession. Second, teacher job satisfaction is the main carrier, i.e. mediator (Hayes, 2013), of all indirect effects present in the model, for within and between estimates. Third, relative differences in classroom discipline carries all indirect effects of demanding student intake characteristics within schools (see Table 2.5). Nevertheless, these were small effects in contrast to the school climate factor estimates. Positive teacher-student relations present additive effects on job satisfaction, and indirect effects on teachers' intentions to leave the profession.

In summary, school climate factors such as teacher-student relations and classroom discipline can moderate individual outcomes because they present contextual effects and/or because they serve are mediators of differences in the composition of the school intake. In addition, teacher job satisfaction plays both roles, carrying contextual effects as well as being the main mediator of all effects.

Limitations and advantages

In spite of the representative sample design of TALIS 2013, the current study has some limitations in terms of generalizability. The sampling frame of the study pertains to teachers from lower secondary levels (7th to 9th grades). Therefore, these results are not necessarily generalizable to primary school teachers in Chile. The second limitation of the current study refers to the nature of the outcome: teacher turnover intentions. These indicators may be dismissed as not objective in nature, in contrast to measures of actual turnover. Nevertheless, as noted earlier, meta-analytic estimates of the relation of turnover intentions with actual turnover are substantial, with r=.45 (Griffeth et al., 2000) after including controls. Studies conducted with nationally representative samples of teachers in the US have also estimated this relationship, yielding positive estimates (Cha, 2008; Finster, 2013). Finster (2013), in particular, found that teacher turnover intentions are highly predictive of teacher retention (teachers who remain in the school), in contrast to leaving the school system (odds ratios of 3.11, r=.42). Moreover, using data from the National Teacher Evaluation System in Chile, which is a large sample (n>6000) yet not representative, Carrasco and Banerjee (2015a) derived estimates of r=.16 (OR=1.20) without controls, and r=.10

(adjusted OR=1.14) with controls. Regardless of its proxy nature, teacher turnover intentions may serve as a sufficient benchmark for human resources practices to capture factors that are consistently associated with actual turnover.

One of the main advantages of this study is its sample design. This enables us to produce population estimates for all lower secondary teachers at the national level. These estimates may serve as a basis for further studies (Kaplan, 2014), and also as realistic baseline levels for induction programmes which aim to reduce turnover and promote retention. The current study provides insight in this line, by unravelling the interrelationship between relevant factors and turnover intentions.

Connection with previous literature

Our results are consistent with previous organizational literature (Podsakoff et al., 2007), and with teacher labour survey literature (Cha, 2008; Finster, 2013): job environments differences are indirectly related to turnover intentions, via job satisfaction. Our results are also consistent with Rivero (2013), who has reported positive effects of school intake (low income, low achieving schools) on turnover decisions (leaving and switching schools). Complementing these findings, the current study shows that school intake effects on teacher turnover intentions are mostly indirect, and mainly explained by classroom discipline. The link between school intake and teacher turnover intentions is therefore better interpreted as a strain factor to the teaching task (Bakker et al., 2007; Quiñones, Van den Broeck, & De Witte, 2012), given its indirect effects via classroom discipline.[8]

Implications

Unravelling relationships between factors is relevant, as some effects manifest as school differences (e.g., tracking between schools, student segregation), whereas other effects may manifest as within-school differences (e.g., within school segregation). Kelly (2004) suggested that teachers may suffer from tracking, that is teachers are assigned to different students and classroom in a systematic way, and thus, teachers from the same school develop different school experiences. Within estimates

[8] It should be noted that we did not identify significant interactions between resource and demand factors, and so these results are not reported in the present report.

are informative in this regard, as these highlight relative differences between teachers from the same school. School intake and classroom discipline factors presented indirect effects at the within level, thus supporting Kelly's proposition. This is relevant for teacher professional development, because if new teachers were systematically assigned to more "difficult classrooms", they can be expected to suffer from higher risks of turnover.

Within school effectiveness research, there is a divide between academic school climate factors (e.g., discipline) and communal school climate factors (e.g., teacher-student relations) (Gill, Ashton, & Algina, 2004). Some authors advocate more attention to academic factors (e.g., Phillips, 1997), deeming communal aspects as irrelevant. Our current results, however, show that both factors present additive effects on teachers' job satisfaction, and thereby indirect effects on turnover intentions. Thus, the dichotomy of preferring one factor over the other is not advised for a teacher retention focus. These dimensions are better understood as complementary, rather than competing explanations for the teaching career.

Further Research

A possible extension of this research to aim for higher generalizability is twofold. This model can be fitted within a multigroup setting (see Stancel-Piqtak & Desa, 2014, for example) with more countries to assess how general the observed patterns are across other national contexts. Additionally, the model can be expanded to a MSEM with multiple indicators (see Marsh et al., 2012; Nagengast & Marsh, 2012), to account for random sampling error as well as measurement error (Lüdtke et al., 2011) in the estimation of contextual effects. This research could be particularly helpful for illuminating the observed contextual effects. The negative contextual effect of school discipline on teacher turnover intentions seems uncontroversial and easy to follow. However, the reverse contextual effect of teachers' job satisfaction and teachers' intentions to leave the profession seems paradoxical. Similar to other reverse contextual effects (e.g., Big Fish Little Pond), this may be explained by a mix of social comparison and selectivity (Blalock, 1984; Burstein, 1980; Nagengast & Marsh, 2012) of teachers and schools. Nevertheless, this is an open research avenue and needs attention in future work.

Teachers' experiences of different school environments can be captured by different school climate factors, such as school discipline and teacher-student relations. However, the appropriate model specification is crucial to unravel the role that these factors play at different levels of inference regarding teacher turnover intentions. Ignoring current methods (e.g., multilevel latent covariate, MSEM), may lead to underestimation of effects, and as consequence the dismissal of school climate factors and job attitudes as relevant factors for public policy on teacher retention. The present work offers clear evidence that the interplay of these factors explains a significant proportion of variance in teachers' turnover intentions.

Figures and Tables Paper 2

Table 2.1 Descriptives, Population estimates

	Mean	SD	N	VIF	Correlation Estimates: Job Satisfaction	Correlation Estimates: Intentions to Leave	Correlation Estimates: Intentions to Quit
School Characteristics							
Intake	.00	1.00	1535	1.30	-.14 ***	.17 ***	.14 ***
Private	.14	.35	1676	2.87	.01	.02	.05
Subsidized	.33	.47	1676	1.73	.00	-.03	.02
Public	.52	.50	1676	-----	-.01	.02	-.05
Location	4.03	1.52	1377	1.63	-.07	-.02	.01
School Size	34.26	28.62	1249	2.09	.02	-.02	-.04
Socio Demographics							
Age	41.34	11.97	1669	6.19	.09 **	-.08 **	-.12 ***
Experience	15.08	11.85	1484	6.26	.06 *	-.06 *	-.13 ***
Sex (Female)	.63	.48	1676	1.05	.06 *	-.11 ***	-.04
Teacher Qualifications							
Teacher Certificate	.86	.35	1634	1.11	.05 *	-.07 **	-.09 *
Ed. No degree	.01	.07	1651	1.02	.02	.00	.02
Ed. Tech. Degree	.18	.38	1651	1.16	.04	-.01	-.09 **
Ed. Univ. degree	.81	.39	1651	-----	-.04	.01	.08 **
Ed. Postgraduate	.00	.07	1651	1.02	-.03	.01	.02
Working Conditions							
Contract Full time	.63	.48	1642	1.39	.03	-.01	-.03
Work in >1 school	.06	.24	1659	1.04	-.11 **	.05	.02
School Climate factors							
School Principal answers							
Distributed Lead.	11.92	2.56	1351	1.51	.02	.01	.03
Instructional Lead.	12.50	1.74	1369	1.35	-.03	.09 *	.10
Mutual Respect	13.93	2.73	1359	1.24	.12 **	-.06	-.08 *
Shortage	2.51	1.00	1335	1.43	-.16 ***	.13 ***	.07
School Delinquency	6.31	2.63	1319	1.41	-.14 **	.09 *	.08 *
Teachers Answers							
Pos. Int. Relations	13.54	2.22	1544	1.29	.38 ***	-.30 ***	-.26 ***
Class. Discipline	10.21	1.96	1180	1.25	.25 ***	-.23 ***	-.25 ***
Job Attitudes							
Job Satisfaction	.00	1.00	1541	3.27		-.72 ***	-.53 ***
Withdrawal cog.	.00	1.00	1541	2.33			.41 ***
Intentions to Quit	1.15	.95	1521	1.68			

Notes: * p<0.05, ** p<0.01, *** p<0.001 SD=Standard Deviation, N=observations, VIF=Variance Inflation Factor. Intake, Job Satisfaction and Withdrawal Cognitions and, are standardize measures, hence its population mean estimate equals zero. All the rest of covariates below 1, express percentages of dummy variables. All parameters are estimated accounting for sample design, and are estimates of population parameters.

Table 2.2 Regressions Estimates with control variables

		Without Job Satisfaction		With Job Satisfaction	
	Job Satisfaction	Intentions to Leave	Intentions to Quit	Intentions to Leave	Intentions to Quit
	B	B	B	B	B
School Characteristics					
Intake	-.02	.11 *	.04	.09 *	.03
Private	-.08	.03	.01	-.02	-.03
Subsidized	.02	-.02	.01	-.01	.02
Public	---	---	---	---	---
Location	-.10 *	.02	.09	-.06	.04
School Size	.13 *	-.04	-.06	.06	.01
Socio Demographics					
Age	.26 **	-.26 **	-.12	-.07	.00
Experience	-.19 *	.25 **	-.01	.12	-.11
Sex (Female)	.07 *	-.08 *	-.04	-.03	-.01
Teacher Qualifications					
Teacher Certificate	-.02	-.01	.02	-.02	.01
Ed. No degree	.01	-.05	.01	-.05 **	.01
Ed. Univ. degree	.03	-.05	-.03	-.03	-.02
Ed. Tech. Degree					
Ed. Postgraduate	-.01	.00	.01	-.01	.00
Working Conditions					
Contract Full time	-.01	.00	.01	-.01	.01
Work in >1 school	-.09 **	.08 *	.03	.01	-.02
School Climate factors					
School Principal answers					
Distributed Lead.	-.04	.02	.02	-.01	.00
Instructional Lead.	-.04	.10 *	.05	.07	.03
Mutual Respect	.04	.03	.00	.06	.02
Shortage	-.13 **	.12 **	.04	.03	-.03
School Delinquency	-.02	-.04	.01	-.05	.00
Teachers Answers					
Pos. Int. Relations	.36 ***	-.27 ***	-.28 ***	.00	-.09 *
Class. Discipline	.20 ***	-.18 ***	-.21 ***	-.03	-.10 **
Proximal mediator					
Job Satisfaction				-.73 ***	-.53 ***
R^2	.29	.20	.20	.56	.40
N	720.00	720.00	714.00	720.00	714.00

Notes: * p<0.05, ** p<0.01, *** p<0.001, B=Standardized Estimates.

Table 2.3 Regressions Estimates without control variables

		Without Job Satisfaction		With Job Satisfaction	
	Job	Intentions to	Intentions to	Intentions to	Intentions to
	B	B	B	B	B
School Characteristics					
Intake	-.02	-.04	.01	-.05	.00
School Climate factors					
Pos. Int. Relations	.36 ***	-.27 ***	-.28 ***	.00	-.09 *
Class. Discipline	.20 ***	-.18 ***	-.21 ***	-.03	-.10 **
Proximal mediator					
Job Satisfaction				-.73 ***	-.53 ***
R^2	.29	.20	.20	.56	.40
N	720.00	720.00	714.00	720.00	714.00

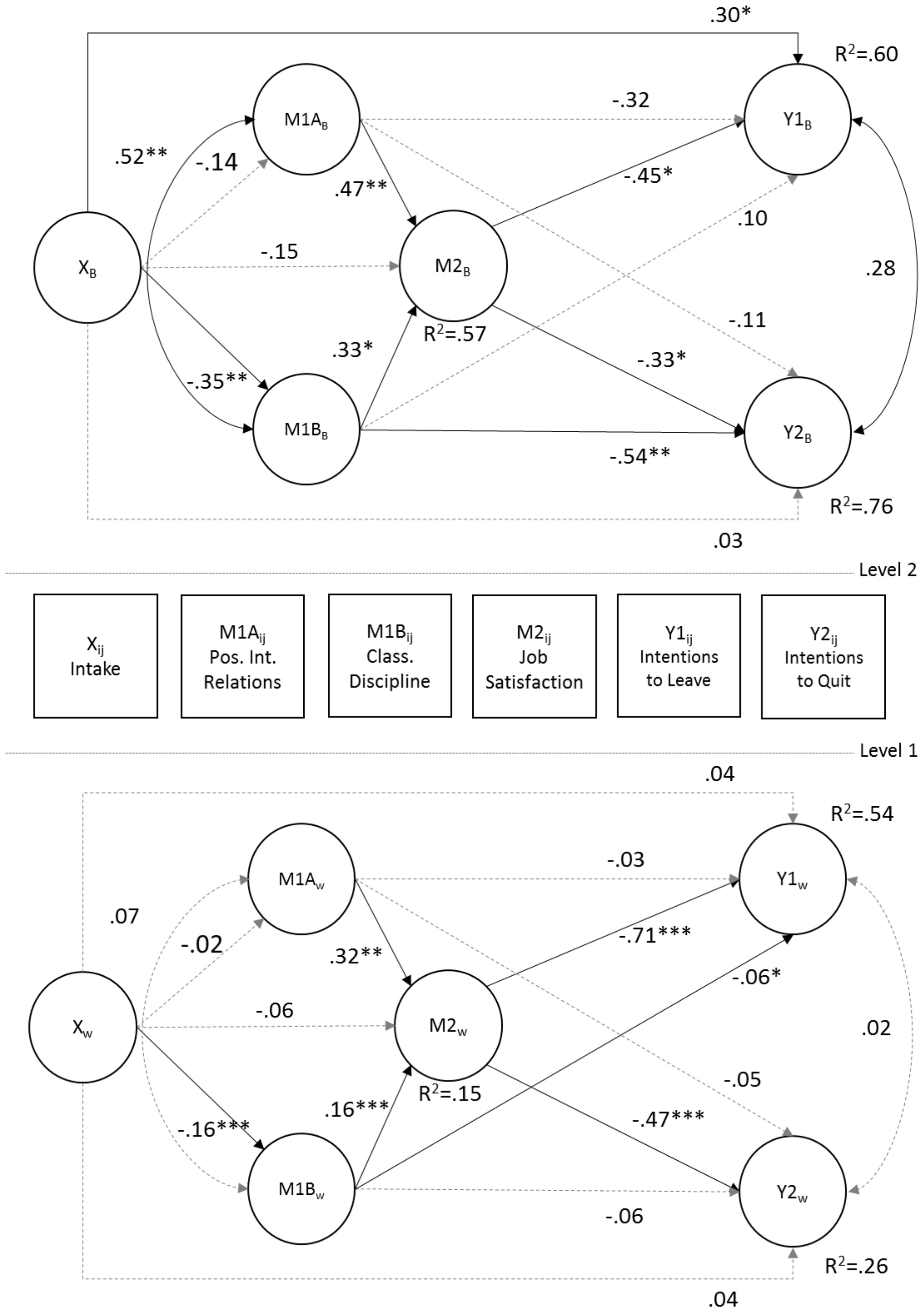
Notes: * p<0.05, ** p<0.01, *** p<0.001, B=Standardized Estimates, N=observations.

Table 2.4 Model Sequence

	Deviance	SCF	Parameters	cd	LRT	df	P
Model 0	27379.88	1.37	23.00				
Model 1	26069.67	1.47	27.00	2.04	641.66	4	.00
Model 2	25848.70	1.46	35.00	1.43	154.09	8	.00
Model 3	25730.26	1.41	45.00	1.21	97.52	10	.00

Note: Deviance= -2*Log Likelihood, SCF=Scaling Correction Factor for -2LL estimates with MLR estimators, Parameters=count of freely estimated parameters by the mdoels, cd = difference test scalling correction factor, LRT=Likelihood Ratio Test, df=degrees of freedom, P= P value.

Figure 2.1 Multilevel SEM diagram, standardized estimates



Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$ Dashed lines are non significant path coeficients, non dashed lines are signficant path coefficientes. Curve lines expresses covariances. R^2=determination coefficient or percentage of predicted values. Intrac Class Correlation values for each variable are under a null model: $Y1_{ij}$ = 8.6%, $Y2_{ij}$ = 10.9%, $M2_{ij}$ = 12.5%, $M1A_{ij}$ = 12.0%, $M1B_{ij}$ = 16.6%, and X_{ij} = 30.7%.

Table 2.5 Indirect Effects, Standardized Estimates

Parameter	Within B	Within (SE)		Between B	Between (SE)	
X⇒M1A⇒M2	-.01	(.01)		-.07	(.06)	
X⇒M1A⇒M2⇒Y1	.00	(.01)		.03	(.03)	
X⇒M1A⇒M2⇒Y2	.00	(.01)		.02	(.02)	
X⇒M1B⇒M2	-.03	(.01)	**	-.12	(.06)	
X⇒M1B⇒M2⇒Y1	.02	(.01)	**	.05	(.04)	
X⇒M1B⇒M2⇒Y2	.01	(.00)	**	.04	(.03)	
X⇒M1A⇒Y1	.00	(.00)		.04	(.05)	
X⇒M1A⇒Y2	.00	(.00)		.02	(.03)	
X⇒M1B⇒Y1	.01	(.01)		-.03	(.09)	
X⇒M1B⇒Y2	.01	(.01)		.19	(.09)	*
M1A⇒M2⇒Y1	-.23	(.02)	***	-.21	(.09)	*
M1A⇒M2⇒Y2	-.15	(.02)	***	-.15	(.09)	
M1B⇒M2⇒Y1	-.11	(.03)	***	-.15	(.10)	
M1B⇒M2⇒Y2	-.08	(.02)	***	-.11	(.07)	

Note: * p<0.05, ** p<0.01, *** p<0.001. B = Estimate, SE=Standard Error, X = School Intake Characteristics, M1A = Teacher-Student Positive Interpersonal Relations, M1B = Classroom Discipline, M2 = Teacher's Job Satisfaction, Y1 = Teacher's Intentions to leave the profession, Y2 = Teachers intentions to quit the school.

Table 2.6 Contextual Effects

Parameter	Within B	Within E	Within (SE)	Between B	Between E	Between (SE)	Contextual B	Contextual E	Contextual (SE)
M1Bij⇒Y2ij	-.06	-.03	-.02	-.54	-.20	-.08 **	-.48	-.17	-.08 *
M2ij⇒Y1ij	-.71	-.72	-.02 ***	-.45	-.36	-.16 *	.26	.36	-.17 *

Note: * p<0.05, ** p<0.01, *** p<0.001. B=Standardized estimates, E = Unstandardized Estimate, SE=Standard Error of E, M1B = Classroom Discipline, M2 = Teacher's Job Satisfaction, Y1 = Teacher's Intentions to leave the profession, Y2 = Teachers intentions to quit the school.

Paper 3. Indirect effects of bullying on school mathematics achievement in Chile

Diego Carrasco
Robin Banerjee

University of Sussex, UK

Reference:

Carrasco, D., & Banerjee, R (2015). Indirect effects of bullying on school mathematics achievement in Chile. *Unpublished manuscript.*

Abstract

Students who experience bullying at school present different negative outcomes, including lower academic achievement. However, the process by which bullying is connected to academic achievement is not clear. Using the Trends in International Mathematics and Science Study (TIMSS) dataset from Chilean schools in 2011, we sought to estimate the indirect effects of bullying on mathematics achievement via two key socio-motivational factors, namely school belonging and students' engagement. Results of our multilevel latent covariate analyses showed that schools' bullying rates were predictive of school differences in mathematics achievement, but that these effects were explained by broader characteristics of the school environment such as perceived levels of safety and discipline. Crucially, the hypothesized indirect pathway was evident at the within-school level, showing that individual experiences of bullying are related to a poorer sense of engagement with the school as a whole, as well as poorer classroom engagement.

Keywords: bullying, belonging, engagement, achievement, multilevel

Indirect effects of bullying on school math achievement in Chile

Violence tends to disproportionally affect youths (Vivolo, Holt, & Massetti, 2011). Elgar and colleagues (Elgar, Craig, Boyce, Morgan, & Vella-Zarb, 2009), using the Health Behavior in School-aged Children (HBSC) survey, estimated rates of self-reported bullying for 37 countries among 11 year old students. The rates vary between approximately 1.8% and 20% (Mean=9%, SD=.1%). Similarly, Contreras and colleagues (Contreras, Elacqua, Martinez, & Miranda, 2015), using secondary data from Trends in International Mathematics and Science Study 2011 (TIMSS 2011), estimated the rates of experiencing being a victim of bullying among 52 countries. Researchers found, for example that among 4th graders, between 3.4% to 48.7% (Mean= 13.1%, SD=8.7%) have experienced at least one form of physical violence in the last month. Thus, victimization is considered a common experience in elementary and secondary school (Jansen et al., 2012).

Chile is a country with a notably high prevalence of bullying. The Global School-Based Health Survey 2004, which surveyed students at ages 13 to 15 years, shows that 47% of Chilean students have been bullied in the past month (Fleming & Jacobsen, 2009). The “Encuesta Nacional de Juventud” from 2012, estimates that about 29% of youth (covering ages 15-29 years) have experience some form of physical or psychological form of violence at school (INJUV, 2013). Chile appears in 6th place for physical violence victimization among 4th grade students in TIMSS 2011, with a rate of 18.4% (Contreras et al., 2015).

Bullying and school adjustment

Experience of school bullying is known to have detrimental effects on pupils’ wellbeing and school adjustment. Depression, anxiety and in worst cases suicide have been linked to experiences of bullying (e.g., Espelage & Holt, 2013; Hertz, Donato, & Wright, 2013). In addition, bullying has been reported to have a negative association with academic achievement. Meta-analytic estimates for the relation between bullying victimization and academic achievement suggest an *r* of -.10 (Nakamoto & Schwartz, 2010). Causal inference estimates, which compared bullied students and non-bullied students, matching students on a range of other characteristics, found differences of 9-

13 points with TIMSS 2011 and PIRLS 2006[9] among the Italian students. These estimates are of similar effect size as class reduction or the improvement of teaching abilities (Ponzo, 2013). Perhaps partly related to these effects, bullying experience also has being associated with lifelong consequences, such as later violence, conviction, drug use and low job status (Farrington & Ttofi, 2011; Ttofi, Farrington, & Lösel, 2012).

Bullying is a cause of concern, not only for victims of bullying, but also for schools as a whole, because bullied students may show counter-violence. In its most extreme form, bullying has been linked to school shootings in the US (Cunningham, 2007). In fact, the Secret Service assert one commonality among school shooters in the US: 71% of them had been targets of bullying (Espelage, Hong, Rao, & Low, 2013). Thus, in broad terms, bullying appears to feed more violence. However, who is the bully and who is the bullied can vary in time within schools (Taki, 2009). Thus bullying is not only an individual experience; it also behaves as a group phenomenon.

Contextual effects of bullying

Bullying behaviour is a social group process highly present in the school environment (Azeredo, Rinaldi, de Moraes, Levy, & Menezes, 2015; Woods & Wolke, 2004), and is understood to display contextual effects. For example, peer group level aggressive behaviour at time 1 moderate individual student aggressive behaviour at time 2, even after controlling for individual differences. Thus, students in more violent peer groups are more likely to display aggressive behaviour (Espelage, Holt, & Henkel, 2003). Classroom bullying levels also moderates the relationship between rejection and victimization for girls. Thus, in classrooms with higher rates of bullying, rejected girls are more likely to be victims of bullying, than in classrooms with lower levels of bullying (Isaacs, Voeten, & Salmivalli, 2013). Indeed, schools with higher levels of bullying rates are expected to have students with lower wellbeing, even after controlling for students’ own experience of bullying (Konu, Lintonen, & Autio, 2002). These negative links suggest that schools with higher levels of bullying most likely offer inferior environments for learning.

[9] Progress in International Reading and Literacy Study (PIRLS), wave 2006.

In fact, the contextual effects of differences in the prevalence of bullying are likely to extend to academic outcomes. Schools with higher bullying rates have been associated with higher dropout rates (e.g., Cornell, Gregory, Huang, & Fan, 2013; Townsend, Flisher, Chikobvu, Lombard, & King, 2008), and school-level regressions with data from Virginia in the US have found negative relations between school prevalence of bullying and school passing rates on an academic achievement test (Lacey & Cornell, 2013). Multilevel estimates of bullying rates reported by school principals in Canada are consistent with the same picture: schools with higher bullying rates yield lower academic results (Konishi et al., 2010). Engel and colleagues (Engel et al., 2009) using secondary data from TIMSS 2007, also found a negative relation between school level bullying and academic achievement across different countries (27 out 49) from Europe (e.g., Hungary, Ukraine), Asia (e.g., Singapore, Hong Kong) and the Middle East (e.g., Israel, Jordan, Lebanon). Román & Murillo (2011), using secondary data from the Second Regional Comparative and Explanatory Study (SERCE), estimate the relation between the experience of bullying and classroom bullying rates across 15 Latin-American countries. The authors found consistent contextual effects of bullying rates across countries. That is, classroom-bullying rates have a consistent negative relation with academic achievement, regardless of students' individual experience of bullying.

Previous literature on school bullying in large scale assessment settings has focused on the size of the relationship between bullying and achievement (Engel et al., 2009; Ponzo, 2013; Román & Murillo, 2011). However, these studies do not answer why bullying is connected to academic achievement. Furthermore, although traditional multilevel models are informative of the effect of covariates at different inferential levels, there are plausible interrelations with the covariates in the model. Thus, these model specifications cannot account for complex relations such as indirect effects. To overcome these limitations in the present study, we fitted a theory driven model, estimated contextual effects (after controlling for a variety of school characteristics) with a multilevel model, and with a multilevel SEM we estimated the indirect effects of bullying on math achievement, via belonging and academic engagement.

Conceptual Model

Buhs and colleagues (Buhs et al., 2006, 2009) propose an indirect effect model for the link between peer rejection and achievement. In this model, the authors stipulate that academic achievement is damaged by peer rejection because of two reasons: students who suffer from peer rejection tend to participate less in classroom activities because of their social exclusion; and students who suffer peer maltreatment will avoid school as whole. The general model can be expressed by the following: peer rejection ➔ chronic peer victimization ➔ lower classroom participation ➔ achievement problems. In this framework, bullying is an example of student victimization. One of the main assumption of this model is that aversive social experiences are stressful for people, who proceed to exhibit social and emotional disengagement from the source of stress (Buhs et al., 2009). Although the original model of Buhs and colleagues has been tested with young school children (pre-primary and primary students) studies with 9th graders support the first path: from bullying to engagement (Mehta, Cornell, Fan, & Gregory, 2013). Thus, bullying is expected to have negative effects on students' school involvement, because of its aversive effects and students' consequent disengagement.

Research on engagement and academic achievement (J.-S. Lee, 2014) asserts that emotional engagement with the school (feelings of school belonging) influences academic achievement indirectly, via behavioural engagement. Lee (2014) supported the engagement hypothesis using secondary data from PISA 2000[10] from the US sample. The implication of this model is twofold: a) more emotionally engaged students are more likely to put greater effort into classroom activities, and thus achieve better results; and conversely b) more emotionally disengaged students are less likely to put higher effort into classroom activities, and thus achieve worse results.

Additionally, the work of Konishi and colleagues (Konishi et al., 2010) suggests that school connectedness may also act as a *buffer* against the negative effects of bullying. School connectedness, school belonging, and school bonding are all different forms of emotional engagement with the school environment (Wormington et al., 2016). The buffer hypothesis states, in positive terms, that even if students suffer from bullying they may get average achievement results to the extent that teachers and peers sustain their sense of school belonging (Norwalk et al., 2015). Moreover, the negative

[10] Program for International Student Assessment 2000

implication of the buffer hypothesis, entails that if students already feel disengaged from the school environment, and also suffer from bullying, more negative results are expected, in contrast to students with average levels of school emotional engagement.

The Present Study

These views of engagement and its expected effect on student withdrawal behaviour because of bullying are complementary. We integrate these previous models under the causal assumption that bullying has negative effects on academic achievement because it negatively affects the engagement process of students with learning. We can express the expected general relations in the following manner: being bullied ➔ lower school belonging ➔ lower classroom engagement ➔ lower achievement (see Figure 3.1). It is not yet clear if the same mechanism can serve to explain variations in achievement within and between schools, but given the previous literature on school-level associations between bullying prevalence and academic achievement, we expect between-school effects as well as within-school effects. In other words, schools with higher levels of bullying are expected to have students with lower levels of school belonging and classroom engagement, and thus present lower academic results. Additionally, to operationalize the buffer hypothesis of school belonging, we include an interaction term between bullying and students sense of belonging. Finally, we include socio demographics (age, sex), and school broader characteristics (e.g., school SES intake, type of school administration and school principal rates of safety, discipline, and academic emphasis) as control variables which have known effects on school achievement (Mullis, Martin, Foy, & Arora, 2012).

To summarise, in the present study, our line of inquiry is twofold. We are interested in the contextual effects of bullying on school outcomes, as well as the motivational route by which bullying is indirectly linked to academic achievement. Our research questions are: a) To what extent school-level variation in academic achievement can be explained by school-level variation in bullying, and if so, is this the case over and above broader school characteristics such as overall levels of discipline, safety and academic emphasis? and b) Are individual, within-school experiences of bullying predictive of academic achievement, and if so, are these explained by feelings of school belonging and engagement?

Figure 3.1: conceptual diagram

Method

Participants and Data sources

We use the national representative sample of eighth graders from Chile in the Trends in International Mathematics and Science Study 2011 (TIMSS 2011). This study is made publicly available by the International Association for the Evaluation of Educational Achievement (IEA) website (http://rms.iea-dpc.org/). The Chilean data comprise a nominal sample of 5835 students, nested in 193 schools, with a mean age of 14.20 (SE=.01) and a balanced gender population (Female=53.01%, SE=.01). This study uses a two stage sample probability design, in which schools are randomly sampled, and intact classrooms are selected at a second stage from the target 8th grade (Joncas & Foy, 2011). Given the nested nature of this data (Snijders & Bosker, 2012), the Chilean representative sample reaches an effective random sample size of approximately 428 students for the Mathematics achievement measure.

Measures

The TIMSS 2011 public data release provides researchers with thorough documentation of the questionnaires and computed scales. Full details of the variables used in the current study can be found on the TIMSS & PIRSL website (http://timssandpirls.bc.edu/). Details of these are also included in Appendix (Paper 3).

Math attainment. TIMSS 2011 used a matrix sampling design in which students answered one of four randomly assigned booklets of 10 to 15 items each (L. Rutkowski et al., 2010). These items covered different mathematics domains such as numbers, algebra, geometry, and data and chance competency. Furthermore, these questions were designed to measure knowledge, application and reasoning. Students' responses were modelled via item response theory (IRT), and proficiency scores were estimated using the MGROUP software (Foy, Brossman, & Galia, 2011) producing 5 plausible values to generate population level proficiency estimates (Von Davier, Gonzalez, & Mislevy, 2009). These scores were set to have a mean of 500 and standard deviation of 100 for all participating countries. All analyses including this outcome were conducted including all five plausible values, using imputation techniques to yield combined estimates (L. Rutkowski et al., 2010).

Socioeconomic status. This was measured by combining students' responses to "Number of Books at home", "Number of study supports" (internet connection, own room, both)", and highest level of education of either parent. These questions were modelled via partial credit IRT, set to a mean of 10 and SD of 2 (Martin, Mullis, Foy, & Arora, 2011), whereby a higher number means a higher socio economic status. In the present study, we used the provided estimates from the TIMSS 2011 data release.

School administration. Three dummy variables were created to identify private school, public schools and subsidized schools. This information was retrieved from the stratification variables from the public data release. We used public schools as reference category.

Gender. This was dummy coded, leaving males as the reference category (female=1, male=0).

Age. Students' age was computed in years, using the test date as a reference. We divided this covariate into three groups: Younger, Expected and Older. We categorized all students according to their implied age of entry at year one of primary school, by March 31 2003. Thus, students who were younger than 6 years at that point were classed as younger, those students who were 6 to 7 were classed as the expected age, and those students who were aged 7 years or more were categorized as older. This last group may correspond to students who may have suffered from previous grade retention.

Engagement. Students indicated their level of agreement, using a Likert type scale of 4 levels, to five different items, such as: "I am interested in what my teacher says", "My teacher gives me interesting things to do", "I think of things not related to the lesson" (reverse coded). These responses were modelled using a partial credit IRT, and were set to an international mean of 10 and standard deviation of 2 (Martin et al., 2011), where higher scores meant more student engagement.

Belonging. Students indicated how much they felt a part of the school, expressing their levels of agreement to three items: "I like being at school", "I feel safe when I am at school", "I feel like I belong at this school", using a four level Likert type scale. We created a mean score with these responses, which yielded an observed alpha of .69.

Bullying. Students indicated the frequency with which they experienced different forms of bullying at school, such as: being called names, being left out, being physically hurt, being forced to do things by other students. In total, this scale included 6 items, to each of which students responded if the event has ever happened, happened a few times a year, once or twice in a month, or at least once a weak. Students responses were combined using a partial credit IRT model, scaling its international mean to 10, and its standard deviation to 2 (Martin et al., 2011), where higher levels meant more experience of bullying by students. This scale was discretized into three ordinal levels of frequency of bullying, at weekly, monthly and almost never bullied. We turned this estimated score into a dummy variable. Those students who suffered from some form of bullying, at least within the last month were flagged with 1, and the remainder of the students were left as the reference category.

School safety. Math teachers from the target grade from each school indicated their level of agreement to five different items referring to school safety and order. This scale included items such as: “This school is located in a safe neighbourhood”, “I feel safe at school”, “The students are respectful of the teachers”. Teachers’ responses were modelled via a partial credit IRT, like the rest of the scales, such that higher numbers in the scale indicated relative safer schools. This covariate is fixed between schools as only one math teacher per school answered this question.

School discipline. School principals’ responses to 11 items were combined using a partial credit IRT model, and set to M=10, SD=2 for the international average (Martin et al., 2011). This scale assessed the extent to which various student behaviours were problematic within the school, including items such as: arriving late at school, vandalism, theft, intimidation and physical violence. Higher values in this scale indicate higher school discipline.

Academic emphasis. School principals’ responses to 5 items were combined via partial credit IRT model with M=10, SD=2 for the international average of participating countries (Martin et al., 2011). This scale assessed the school emphasis on academic success. Exemplary items of this scale are: “Teachers’ expectations for student achievement”, “Parental support for student achievement”, “Students’ desire to do well in school”. Higher values in this scale indicate higher academic emphasis at the school level.

Analysis Strategy

TIMSS 2011 uses a complex sample design (Heeringa, West, & Berglund, 2009), in which schools are randomly selected, and in a second stage, intact classroom are selected, with an unequal probability. It is a requirement to account for its sampling design, to produce population estimates. Furthermore, this study uses the plausible values method (Von Davier et al., 2009), which combine estimates via imputation procedures, across all available plausible values (L. Rutkowski et al., 2010). We use appropriate variance estimations via fixed and/or mixed effects methods, to account for plausible values and sampling design.

We first estimated the means, standard deviations and standard errors for our selected variables, as well as the effect size, expressed as *r* coefficient, for the relation between each covariate and Math attainment (see Table 3.1). We use a jackknife variance estimation to get valid population estimates via the PV module (K. Macdonald, 2014), and the SVR module (Winter, 2008) in STATA (StataCorp, 2013) in order to estimate these parameters. This procedure replicates official release results for the Chilean sample (Agencia de Calidad de la Educación, 2011). Additionally, to assess plausible sources of multi-collinearity among the selected covariates, we estimated the variance inflation factor for all variables. All covariates yielded a VIF<10, which is the threshold for concern (O'Brien, 2007).

To estimate the relative contribution of each factor onto math attainment, we fit a series of multilevel models using MPLUS v7 (L. K. Muthén & Muthén, 2012) to account for sampling design and plausible values, while scaling weights to effective sample size (Asparouhov, 2006; Rabe-Hesketh & Skrondal, 2006; Snijders & Bosker, 2012). This scaling weight method was preferred over other weight normalization methods, because it produces unbiased estimates of variances (Stapleton, 2002). These results are presented in Table 3.2.

Underlying the relative contribution of all factors to math attainment, there are several interrelationships. To properly estimate the indirect effects and moderation of indirect effects, we fit a multilevel structural equation model (K. J. Preacher et al., 2011a). Within this model, we fit a multilevel mediation with fixed slopes, which resemble the generic model I in Preacher, Zhang, and Zyphur (2011b). Additionally, we included a moderated mediation, in which the indirect effect is moderated by levels of a mediator (type Model 1, in Preacher, Rucker, & Hayes, 2007). In this model, we multiplied the bullying and belonging variables, centred to the grand mean, to get the best estimates of the moderation at the between level (Ryu, 2015). We again used MPLUS v7 to fit this model, including sampling weights, using the same scaling method as before, accounting for plausible values and sample pseudo stratification. We used this third approach to estimate the indirect effects of bullying onto math achievement, via belonging and engagement, jointly at within and between parts of the model.

As noted earlier, the relationship between the experience of bullying and academic achievement, is expected to be mediated by socio-motivational processes. We expected indirect effects on achievement, via the emotive side of motivation (belonging) and its attitudinal aspect in the school setting (engagement in math lessons). Additionally, we tested if belonging serves as a buffer for the negative effects of bullying on achievement. To assess if the indirect effects vary at different levels of the moderator, we calculated a Linear Moderated Mediation Index, LMM for short (Hayes, 2015), at each level.

We relied on a multilevel latent covariate model (Lüdtke et al., 2008), which enabled us to test if these indirect effects are present within or between schools, or both, after inclusion of relevant covariates. The fitted model is presented in Figure 3.2; we follow Ryu's (2015) style of diagram, and depict latent variables as circles, at each level of the model, and use squares for manifest variables.

Figure 3.2: Multilevel SEM diagram

Results

Descriptives

Chile obtained a mean of 416 (SE=2.59) in Math attainment for the population of 8th graders, reaching the 31st position of 42 participating countries. The zero order correlation between students' socioeconomic status and attainment results in an *r* of .4, which is quite high in comparison to other countries in similar studies (Sirin, 2005).

Similar effect sizes can be seen for other school factors, such as school administration (public r=-.34), school safety (r=.32) and academic emphasis (r=.34) when no other factors are taken into account. These results clarify the large differences between schools as educational environments in the Chilean case.

Table 3.1: Population descriptive estimates

Multilevel estimates

Model sequence. The full model sequence in Table 3.2 shows the relative contribution of each factor to explain math achievement. The null model, presents an intra class correlation of .43 (see Table 3.3), indicating that a large amount of the variance of Math achievement is located between schools. Socioeconomic status of the students and the school mean of socioeconomic status, accounts for 1% and 60% for the variation of Math achievement, at the within and between levels, respectively (see M1 in Table 3.3). The socio-demographic factors of sex and age of students were included as controls and accounted for 9% of the within variance, and 2% for the between variance of the outcome (M2). The contribution of students' experience of bullying, belonging and engagement, altogether accounted for 2% and 10% of the variance in achievement within and between schools respectively, after the effects of socioeconomic status, sex and age of students have been controlled for (M5). Finally, the inclusion of the school fixed factors of school safety, school discipline, and school academic emphasis account for an additional 6% of between variance (M6). Overall, the full model accounts 13% for students' variance and 79% of the variance between schools. Each block of variables included in the model sequence improved the model fit at each step (see the LRT results in Table 3.3, comparing each model with the previous model in the sequence).

Main effects. Students' SES was positively related to academic achievement. For every 1 point more above the mean of SES within schools, students were expected to score 4.55 points (SE=.94, p<.01) higher on math achievement in contrast to their schoolmates. Similarly, schools with 1 more point above the SES mean presented 23.46

points (SE=4.27, p<.01) above the estimated grand mean for achievement. The contextual effect of this factor is therefore estimated at 18.91 points (SE=4.29, p<.01). Thus, the expected difference in math achievement between two students with similar SES characteristics, who attend schools differing by 1 unit on the school SES mean, would differ on 18.91 points in mathematics achievement, which is .24 standard deviations of the outcome (b/ySD)[11]. Additionally, private schools achieve 44.33 points (SE=13.13, p<.001) more over the grand mean of schools, even when all other factors are controlled for.

Socio-demographics factors were of great relevance. Female students were expected to achieve 22.39 points (SE=2.80, p<.01) less than their male counterparts, while in terms of age, only the ‘older’ students differed from the rest, scoring about 38.36 points lower (SE=3.35, p<.01) than the students with the ‘expected’ age. Because these factors were entered as controlled variables, these are centred to the grand mean. Thus, all effects are estimated regardless of school composition in terms of sex and age (Heck & Thomas, 2015).

In model 5, bullying, belonging and students’ engagement presented interesting effects, before we controlled for other school factors. In this model, bullying and belonging showed larger effects between schools, than within schools. Thus, this entails significant contextual effects: students who attended schools where 100% of the students reported bullying could be expected to achieve 48.55 points lower (SE=19.13, p<.01) than schools at the grand mean of bullying. If we change the metric of the covariate to increments of 10%, it means that for every 10% increase in the bullying rate at the school level, schools may be expected to achieve 4.85 points lower than average schools. With respect to school belonging, students who attended schools with 1 point over the grand mean of school belonging would be expected to achieve 43.75 points (SE=11.55, p<.01) more on mathematics. If students’ characteristics are held constant, these are the expected differences attributable to the school environments. In contrast, students’ classroom engagement did not present larger between effects over within effects. More engaged students present higher scores in math achievement in comparison with their peers (b=5.88, SE=.90, p<.01, see M5 in Table 3.2).

However, when school safety, school discipline and school academic emphasis were controlled for, these contextual effects no longer reached statistical significance

[11] Unstandardized estimated, divided by the residual variance of the outcome.

(see M6 in Table 3.2). For each unit of change of School Academic Emphasis, schools reach 3.97 points over the grand mean (SE=1.25, p<.01). Similarly, schools higher in School Safety could be expected to have 2,77 points more, for every unit of change (SE=1.27, p=.03), although differences in school discipline did not explain much variance in achievement (b=.79, SE=1.56, p=.62). In the final model, then, bullying, belonging and engagement did not present significant effects between schools because the previously observed effects had been washed out following inclusion of schools' levels of safety and academic emphasis. At the within level, however, differences between students' engagement within schools did have an effect on the outcome. For 1 unit of change in engagement, students achieved 6.01 points more (SE=.89, p<.01) than their schoolmates.

Multilevel Latent Covariate Estimates/Multilevel SEM

Overall fit. Because MPLUS estimates one model for each of the plausible values, each fit index has a point estimate and a standard deviation. The estimated model presents a good fit: $\chi^2(4) = 13.85$ (SD=.92), p= .01, CFI=.99 (SD=.00), RMSEA=.02 (SD=.00), SRMR within = .01 (SD=.00), SRMR between = .02 (SD=.00). In Tables 3.4 and 3.5, we report the main parameters of the model, as well as the indirect effect estimates. In the Appendix (Paper 3) material, in Tables 3.6 and 3.7, we present the rest of the parameters of the model, including the estimates for the controlled variables.

Main effects on math achievement. Relative differences between students' SES within school were positively related to math achievement (b=4.62, SE=.95, p<.01). Likewise, schools with an intake showing a higher mean SES also achieved higher scores (b=26.82, SE=5.00, p<.01). This factor, presents a contextual effect of 22.20 points of difference in mathematics achievement for every point increase of school mean SES (SE=5.01, p<.01). Additionally, other school factors accounted for overall math achievement. Private schools reach 33.50 points more over the grand mean (SE=14.80, p=.02). Schools with higher academic emphasis also could be expected to have higher scores (b=3.58, SE=1.33, p<.01). No other school level factors present significant effects.

Students' sex and age explained some of the differences in math achievement. In general, female students presented lower scores than male students (b=-22.44, SE=2.85, p<.01). Students older than expected also presented lower scores than the rest of the population (b=-38.84, SE=3.27, p<.01).

Indirect effects. We decomposed the total and indirect effects of Bullying on Math Achievement. These effects were significant only at the within part of the model. There were overall negative indirect pathways from bullying to achievement (b=-2.07, SE=.52, p<.01). These indirect effects were carried via our more proximal mediator, namely students' engagement (b=-1.15, SE=.34, p<.01). Additionally, there was a negative indirect effect via belonging which itself is related to mathematics achievement via classroom engagement (b=-.98, SE=.21, p<.01). Finally, the indirect effect via classroom engagement was conditioned by students' belonging levels (LMM=.08, SE=.02, p<.01), such that the indirect effect was larger when students had lower levels of belonging (-1SD, b=-1.29, SE=.29, p<.01) and smaller when students had higher levels of belonging (+1SD, b=-.68, SE=.15, p<.01). Thus, students' sense of belonging with the school appeared to buffer the negative effects of bullying experience.

Discussion

Our first research question referred to the contextual effect of bullying on academic achievement. In the present study, we found evidence in favour of such a contextual effect. Specifically, schools with higher levels of bullying presented lower math achievement, regardless of students' individual experience of bullying. However, when we controlled for school safety, discipline and academic emphasis, this contextual effect was no longer significant. Thus, the contextual effects of bullying on math achievement seem to be explained by other observed characteristics of the school environment.

Our second line of inquiry concerns the indirect effects of bullying on achievement, via socio-motivational processes. Our results support this assumption: students who experience bullying showed lower math achievement, but this link is explained by students' engagement. Specifically, students who suffer bullying were less engaged with lesson, this in turn was predicted by lower levels of engagement/belonging with respect to the school as a whole. Moreover, the indirect pathway via classroom engagement was also *moderated* by the level of students' sense of school belonging: students with higher belonging presented a smaller indirect effect of bullying on achievement, in comparison to students on average levels of belonging.

Our results are consistent with previous findings. Bullying does appear to have negative contextual effects on achievement (Engel et al., 2009; Konishi et al., 2010; Román & Murillo, 2011), but these effects cease to appear once other school differences are accounted for. Academic emphasis, safety, and orderly environment are known effectiveness factor for school achievement (Mullis et al., 2012). Bullying event rates may have indirect effects on school academic achievement, via their relationship with school safety and discipline, as a bottom-up effect (K. J. Preacher et al., 2010). We did not test such model specifications regarding the interplay with covariates, but this is a line of research which would serve to illustrate how bullying contextual effects may work, via other observed school characteristics. Nonetheless, we can conclude from the present analyses that school-level variations in bullying serve as a meaningful predictor of school differences in achievement only to the extent that the reflect variations in broader school characteristics regarding levels of safety and discipline.

In contrast, the within-school link between individual experiences of bullying and achievement seem to be explained by its relation to socio-motivational processes. Our results are consistent with previous authors' models and suggestions (Buhs et al., 2006, 2009; Mehta et al., 2013; Wormington et al., 2016), in which bullying is depicted as a stressful factor that diminishes school belonging, which in turn diminishes students' engagement with school lessons and thus is negatively related to academic achievement. Furthermore, we find positive evidence for the buffer effect of school belonging (Konishi et al., 2010), even with a different set of measures. Konishi and colleagues (2010) assessed school engagement via students' and teachers' relationships, whereas in the present study we use direct measure of students' sense of belonging with the school. But our results are consistent in showing that those schools which manage to promote higher levels of school belonging for all their students may also manage to diminish the negative impact that bullying has on individual students' motivation and therefore achievement.

This study has some limitations due to its design, and due to its model specification. First, because of its design, these results relate only to the population of students at grade 8th from Chile. Thus, these results may not generalize to other contexts and other age cohorts. Nevertheless, Wormington and colleagues (2016) compared a similar model in which school belonging was the mediator of the effects of peer victimization and school achievement in two different samples from middle school and high school, and found no large differences for these two age groups. In order to robustly test the proposed model on different ages and different context, this same model can be fitted across all countries that participated in TIMSS 2011, and similarly, can be fitted onto the 4th grade samples. This future work would provide further tests regarding the generalizability of the proposed model.

Because of the current model specification, we have paid no attention to gender differences for this model. There is a substantial literature within the bullying and peer victimization research indicating that gender plays an important role in this phenomenon. For example, boys tend to be more involved in physical bullying, whereas girls tend to be more involve in relational bullying (Berger, Rodkin, & Karimpou, 2008). Moreover, peer rejection is more predictive of victimization for girls, but not for boys, in schools with higher bullying (Isaacs et al., 2013), and so some contextual effects are already known to be different for boys and girls. Females students also seem to be more at risk of suicide related to bullying (Bauman, Toomey, & Walker, 2013; Hertz et al., 2013). These are just a few examples of gender differences and bullying dynamics. In future research, we aim to separate physical and relational bullying, similarly to Contreras et al (2015), in order to estimate if the gap between each type of bullying is similar or not for both genders.

Studies targeting bullying as an outcome frequently inform efforts to prevent this behaviour and also to identify more at-risk populations. The present work provides information regarding the psychological mechanisms that serve as plausible causal links between the experience of bullying and poorer academic achievement. By means of understanding how aversive events such as bullying are linked to different school adjustment outcome, interventions can target the specific mechanism by which bullying exerts its influence on school outcomes (Wormington et al., 2016). This is informative, not only for prevention, but for interventions as well. An example of this line of research was conducted by Norwalk and colleagues (Norwalk et al., 2015). In this study, school belonging was indirectly related to teacher attunement with respect to student victimization. This is a specific type of teacher responsiveness, and refers to how much teachers know about their students' bullying dynamics (e.g., who is being bullied in different ways). Schools with higher attunement also tend to have students with less tolerance for bullying, thus creating a protective peer environment which in turn explains school belonging. Thus, while the present research shows that a broad measure of students' emotional engagement with the school context (feeling that they belonged, feeling safe, and liking school) played a role in the effects of bullying, much more needs to be done to untangle the variables that give rise to this sense of belonging in the first place.

In summary, our work sets an agenda for research on school bullying as a pervasive phenomenon across schools. We have demonstrated that most of its relations with academic achievement are of an indirect nature. Thus, univariate approaches and conventional model specifications may not adequately capture its influence on the outcomes under study, by either ignoring its contextual routes of influence, or by centering attention merely on "significant [main] effects", or by ignoring its relations to other variables. Careful consideration of model specification is essential in this regard, as subtle changes in how to include covariates can lead to different conclusions based on the same data. Our results point to the need to account not only for contextual covariates in order to understand school-level variations, but also indirect and interaction effects involving socio-motivational factors in order to understand within-school effects of bullying on students' achievement.

Figures and Tables Paper 3

Figure 3.1 Conceptual model

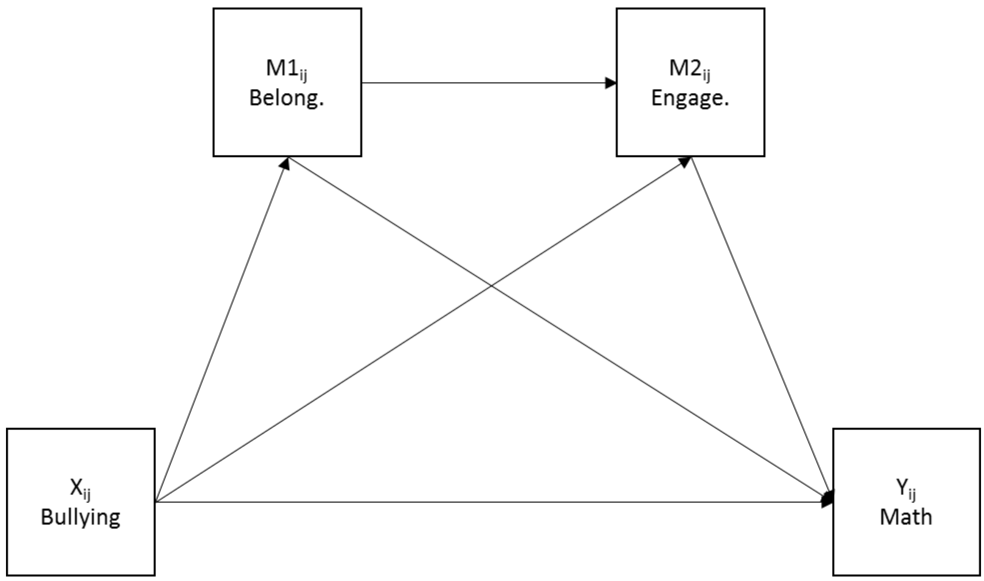


Figure 3.2 Multilevel SEM diagram

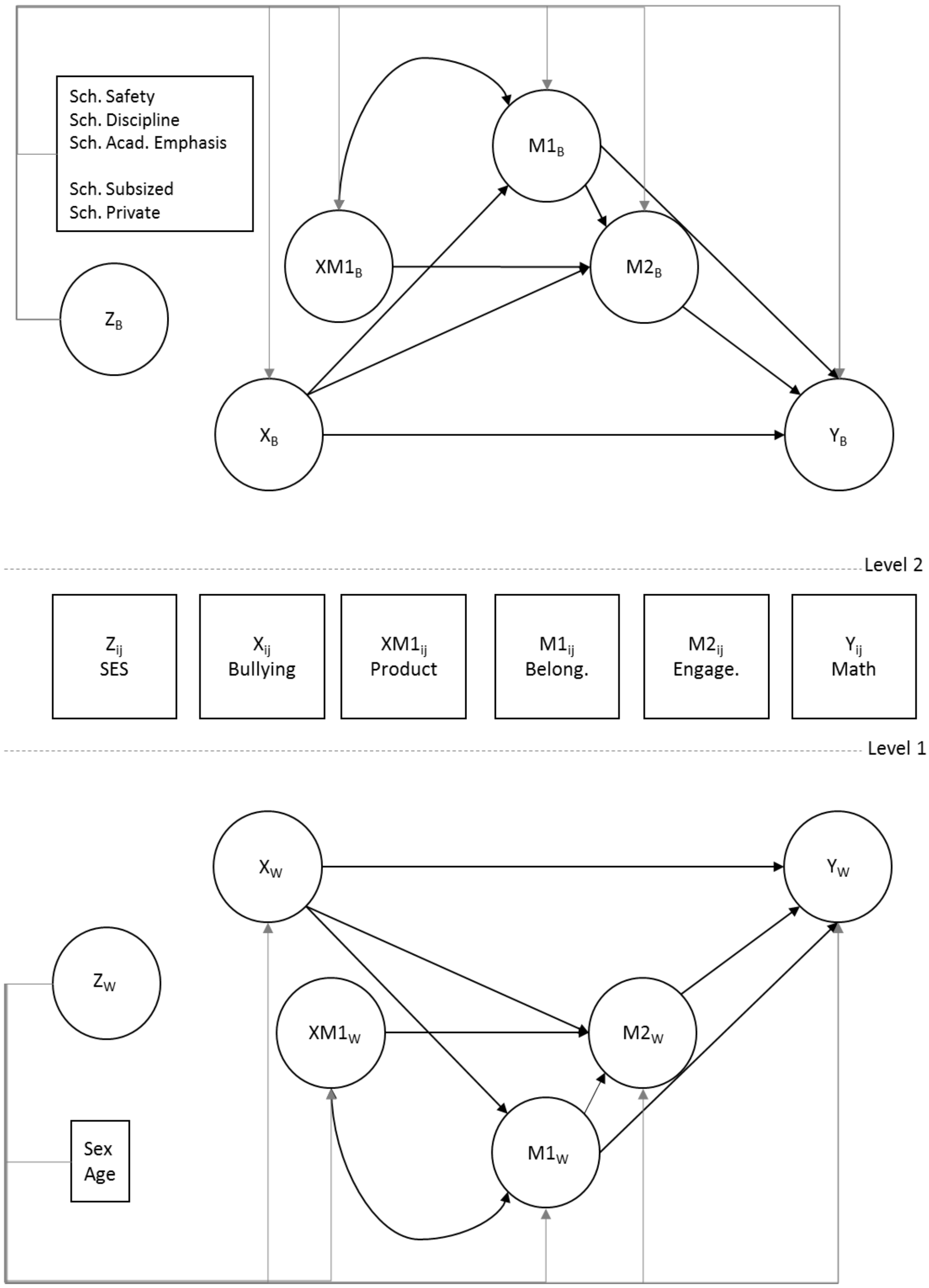


Table 3.1 Population descriptive estimates

	M	SD	SE	VIF	r
Outcome					
Math attainment	416.27	79.65	2.59	1.60	
Socio Demographics					
Socioeconomic Status	9.71	1.64	.05	1.38	.40
Gender					
Female[d]	.53	.50	.01	1.07	-.09
Male[d]	.47	.50	.01	----	.09
Age	14.21	.63	.01		.05
Younger[d]	.13	.34	.01	1.04	.01
Expected[d]	.71	.45	.01	----	.20
Older[d]	.16	.37	.01	1.14	-.26
Students Factors					
Engagement	9.86	1.77	.05	1.16	.11
Belonging	3.38	.64	.02	1.18	.10
Bullying	4.11	1.70	.03	1.08	-.10
Bullying[d]	.38	.48	.01	----	-.10
School Factors					
School Administration					
Public[d]	.46	.50	.01	----	-.34
Subsidized[d]	.47	.50	.01	1.36	.15
Private[d]	.07	.26	.00	1.59	.36
School Safety	9.44	2.35	.18	1.36	.32
School Discipline	9.65	1.78	.15	1.61	.26
Academic Emphasis	8.74	2.37	.17	1.68	.34

Note: covariates flag with [d] are dummy variables, hence their mean estimates are equivalent to percentages estimates. M=Mean, SD= Standard Deviation, SE=standard error, VIF=variance inflation factor, r=Pearson correlation, “----“ are reference categories for dummy variables. Variance Inflation Factor was estimated using only the first plausible value.

Table 3.2 Multilevel fixed effects estimates explaining Math Attainment

		M0	M1	M2	M3	M4	M5	M6
	Intercept	405.52 ***	410.51 ***	411.23 ***	412.33 ***	412.51 ***	412.05 ***	412.00 ***
Between School Factors								
	School mean SES		27.18 ***	24.95 ***	23.11 ***	26.10 ***	26.85 ***	23.46 ***
School Administration	Public (reference)		-----	-----	-----	-----	-----	-----
	Subsidized		12.88	14.13 *	12.54	12.23 *	12.80 *	7.19
	Private		49.40 **	56.06 **	48.92 **	44.48 **	45.52 ***	44.33 **
School Factors	Bullying				-68.87 **	-48.94 *	-48.95 *	-34.18
	Belonging					50.72 ***	42.40 ***	23.15
	Engagement						6.97	7.42
	School Safety							2.77 *
	School Discipline							.79
	Academic Emphasis							3.97 **
Within School Factors								
SES	Student SES		5.49 ***	4.34 ***	4.33 ***	4.33 ***	4.05 ***	4.55 ***
Demographics	Sex							
	Female			-23.41 ***	-23.75 ***	-24.31 ***	-22.96 ***	-22.39 ***
	Male			-----	-----	-----	-----	-----
	Age							
	Younger			.80	.85	.80	.14	-.41
	Normal			-----	-----	-----	-----	-----
	Older			-38.91 ***	-38.35 ***	-37.48 ***	-36.88 ***	-38.36 ***
School Experience	Bullying				-2.00	-1.35	-.40	-.56
	Belonging					3.67	-1.35	-.36
	Engagement						5.88 ***	6.01 ***
Contextual effects								
	SES		21.69 **	20.61 **	18.78 **	21.77 ***	22.80 ***	18.91 ***
	Bullying				-66.87 **	-47.59 *	-48.55 *	-33.62
	Belonging					47.05 **	43.75 ***	23.51
	Engagement						1.09	1.40

Note * p<.05; ** p<.01; *** p<.001.

Table 3.3 Multilevel random effects and fit indices

	M0	M1	M2	M3	M4	M5	M6
Variances							
Within	3659.73 ***	3611.80 ***	3327.53 ***	3325.22 ***	3320.14 ***	3238.63 ***	3206.64 ***
Between	2785.63 ***	1071.51 ***	929.49 ***	809.11 ***	692.42 ***	686.79 ***	549.30 ***
ICC	.43	.23	.22	.20	.17	.17	.14
Model Fit							
Deviance							
-2LL (mean)	65093.92	64531.99	64045.41	63546.03	63493.45	63057.17	53359.61
(SD)	120.45	135.73	150.17	143.26	145.64	126.91	98.77
df	3.00	7.00	10.00	12.00	14.00	16.00	19.00
LRT		.00	.00	.00	.00	.00	.00
Pseudo R square							
R^2 Within		.01	.10	.10	.10	.12	.13
R^2 Between		.60	.62	.67	.72	.73	.79

Note * p<.05; ** p<.01; *** p<.001.

Table 3.4 MSEM main estimates, unstandardized

Parameter			Within E	Within (SE)		Between E	Between (SE)	
Engagement	⇒	Math	6.01	(.91)	***	5.84	(7.23)	
Bullying	⇒	Math	-.55	(2.24)		-63.88	(45.68)	
Belonging	⇒	Math	-.32	(2.63)		41.01	(22.21)	
SES	⇒	Math	4.62	(.95)	***	26.82	(5.00)	***
Belonging	⇒	Engagement	.89	(.05)	***	5.63	(5.65)	
Bullying * Belonging	⇒	Engagement	-.43	(.11)	***	-38.08	(51.73)	
Bullying	⇒	Engagement	-.19	(.05)	***	1.13	(1.73)	
SES	⇒	Engagement	.04	(.02)		-.29	(.52)	
Bullying	⇒	Belonging	-.18	(.02)	***	-.18	(.26)	
SES	⇒	Belonging	-.01	(.01)		-.09	(.03)	**
SES	⇒	Bullying	-.01	(.01)		-.03	(.02)	
Belonging	⇐ ⇒	Bullying * Belonging	.03	(.01)	**	.00	(.00)	*
SES	⇒	Bullying * Belonging	-.01	(.01)		-.02	(.01)	

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, ⇐ ⇒ expresses a covariance estimation.

Table 3.5 MSEM Indirect effects

Indirect Effect Decomposition		Within E	Within (SE)		Between E	Between (SE)
Bullying	Total	-2.62	(2.23)		-70.54	(49.67)
	Indirect	-2.07	(.52)	***	-6.66	(14.41)
	Engagement	-1.15	(.34)	**	6.75	(13.61)
	Belonging	.06	(.48)		-7.35	(11.46)
	Belonging ⇒ Engagement	-.98	(.21)	***	-6.06	(13.56)
	Direct	-.55	(2.24)		-63.88	(45.68)
	LMM	.48	(.17)	**	41.28	(100.39)
Belonging	Total	5.03	(2.30)	*	74.79	(52.02)
	Engagement	5.35	(.90)	***	33.78	(55.98)
	Direct	-.32	(2.63)		41.01	(22.21)

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, ⇐ ⇒ expresses a covariance estimation.

Paper 4. Walk the talk! Civic knowledge, democratic values, and the role of school climate for open discussion. A moderated mediation approach

Diego Carrasco
Robin Banerjee

University of Sussex, UK

References:

Carrasco, D., & Banerjee, R (2015). Walk the talk! Civic knowledge, democratic values, and the role of school climate for open discussion. A moderated mediation approach. *Unpublished manuscript.*

Carrasco, D., & Banerjee, R. (2015). Walk the talk! Civic knowledge, democratic values, and the role of school climate for open discussion: A moderated mediation approach. In IEA International Research Conference 2015. Cape Town, South Africa. *Conference Paper*.

Abstract

The aim of this study was to estimate the association between civic knowledge and democratic values, the mediating factors involved in this effect, and the role of school climate. We evaluated a moderated mediation model in which civic knowledge indirectly explains democratic values via authoritarianism. Furthermore, we specified open classroom discussion as a moderator of this relationship. Results from the Latin-American and Asian samples of the International Civic and Citizenship Study (ICCS, 2009) study partially supported the model. For all tested samples, civic knowledge was positively associated with democratic values, indirectly via inverse relations with authoritarianism. Furthermore, we found evidence of a moderated indirect effect, whereby the indirect effect was larger when students were exposed to more open political discussion in the classroom. This was the case for 7 out of 11 country samples, mainly in samples with a lower mean of civic knowledge or a higher mean of authoritarianism at the country level. Our models fit better for predicting democratic values of gender equality and permissibility of corruption, in comparison with support for ethnic minority rights and support for freedom of speech.

Key words: authoritarianism, democratic values, open climate, indirect, moderated mediation

Walk the talk! Civic knowledge, democratic values, and the role of school climate for open discussion: A moderated mediation approach

The main aim of civic education is the promotion of democratic values, by fostering civic knowledge and the endorsement of democratic attitudes (Lenzi, Vieno, Sharkey, Pastore, & Santinello, 2014). The interpretation of democracy as "a mode of associated living" (Dewey, 1966) requires citizens to behave socially in different contexts. However, the presence of injustice in its various forms erodes the legitimacy of democratic institutions. Prejudice, corruption, and a lack of commitment to equality are primary concerns in this regard.

Past research attests to positive links between civic knowledge and tolerance towards minorities (Caro & Schulz, 2012) and egalitarian values (Lauglo, 2013), and, generally speaking, civic knowledge has positive relationships with several democratic attitudes (Schulz et al., 2010). However, much of this research has been conducted within a univariate framework; that is, estimating the relationships between factors and a single outcome. Consequently, the interrelations between predictors is often ignored. Yet, some of these relationships could be interlinked in complex ways. Of particular interest in the present study is the way in which a number of theoretical frameworks – such as the Dual Process motivation model (Duckitt, Wagner, du Plessis, & Birum, 2002), System Justification Theory (Jost, Banaji, & Nosek, 2004), and the Group-focused Enmity model (Zick et al., 2008) – posit authoritarianism as a precedent for generalized prejudice.

Racism, sexism, and anti-immigrant attitudes are all examples of different forms of prejudiced attitudes. Because attitudes are developed and learned, it is generally thought that these can be unlearned as well (Zick, Küpper, Hövermann, German, & Fenn, 2011). Thus, we aim to integrate the learning hypothesis (Caro & Schulz, 2012) to general theories of social psychology, to offer a plausible explanation of the relationship between civic knowledge and democratic attitudes. Specifically, we address the need to estimate the relationship between civic education and democratic values, explore its psychosocial mediating factors, and consider the role of school climate factors in these relationships. Thus, our research questions are: a) what is the relationships between civic knowledge and different democratic values?, b) why these factors are linked?, and c) what is the role of school climate in these relations?

Authoritarianism and democratic values

Authoritarianism and social dominance orientation are fundamental worldviews (Duckitt & Sibley, 2009) that shape generalized prejudice (Duckitt et al., 2002), sexism (Christopher & Wojda, 2008), and general support for human rights (McFarland, 2010). These two dimensions tap what historically has been called the 'authoritarian personality' (Sibley & Duckitt, 2008). Longitudinal studies stress that these dimensions are consistent precursors of generalized prejudice (Asbrock et al., 2010).

Classically, this anti-democratic syndrome is thought to be more prevalent in individuals of low educational attainment (Lipset, 1959). A recent review supports this (Napier & Jost, 2008): economically disadvantaged groups present higher levels of authoritarianism, especially for the facets of obedience to authority and cynicism (e.g., lack of political interest). According to Houtman (2003), this is due to a limited cultural capital. Low educational attainment and its restricted cultural experience inhibit socio economic disadvantaged groups to expand their understanding of different groups and ideas (Carvacho et al., 2013).

Democratic values and educational attainment

The 'sophistication hypothesis' (Highton, 2009; Luskin, 1990) suggests that people develop the necessary cognitive skills for democracy through education (Evans & Rose, 2007, 2012). Complementing this, socialization theory suggests that the schooling process promotes scientific and democratic values endorsement. Because these are contrary to most prejudice beliefs, the longer individuals are exposed to the educational system, the less prejudiced they are expected to be (Coenders & Scheepers, 2003). However, longitudinal studies have found conflicting results. Studies of the educational effect using the Swiss Household panel (Lancee & Sarrasin, 2015) found that attitudes towards immigrants are mainly present between individuals, and these differences remain stable as individuals pass through education. Another panel study, conducted in Belgium found that ethnocentrism can even grow over time (Hooghe, Meeusen, & Quintelier, 2013). Specifically, while students from the general academic track developed more tolerant attitudes, students in lower tracks, such as technical and vocational education, developed less tolerant attitudes. A panel study conducted in the Netherlands (Vollebergh, 1996), following students between 12 and 18 years approximately, in waves between 1986 and 1990, and from different educational streams (Lower Technical, Lower Secondary, Higher Secondary, and Pre-university Education), found differences between and within students. There were differences between students depending on the educational track, these differences were larger after 4 years, and students in the more academic tracks diminish their authoritarianism endorsement after 4 years. Thus, longitudinal research shows that effects of education are complex and not uniform across individuals and educational systems.

The moral enlightenment hypothesis (Easterbrook, Kuppens, & Manstead, 2015) of the educational effect onto democratic attitudes has two requirements. Firstly, it must provide individuals with increased and more sophisticated knowledge; otherwise, mere school attendance would have no mechanism to countervail prejudiced beliefs. Secondly, the schooling process must promote democratic values endorsement; otherwise, the promotion of different values would drive social attitudes different from the expected outcome.

The first assumption makes a crucial distinction for understanding the education effect on democratic attitudes, by dissociating effective knowledge acquisition and years of schooling. This, for example, may explain why students in the lower tracks of the educational stream, as in the Belgian sample (Hooghe et al., 2013), develop less tolerant attitudes, while the students in the academic track do grow in more tolerant attitudes. This distinction alone cannot explain the observed differences in the Swiss Household Panel Survey (Lancee & Sarrasin, 2015), in which even more educated individuals in some cases may also endorse less tolerant attitudes over time.

For this matter, the second assumption is key: the educational effect occurs inasmuch as it promote certain values and not others. Within civic and citizenship education, it has been noted that what schools do to promote citizenship occurs via certain conceptions of being a “good citizen”. An analysis of ten citizenship programmes in the US showed that some programmes’ emphasis lies in character and personal responsibility, whereas others emphasise community service and participation, or social critique and social justice (Westheimer & Kahne, 2004). Acording to the authors, programmes which emphasize the role of participation in the community showed positive influence on leadership efficacy and social capital for community development, but it did not change students’ explanations of poverty or their interest in politics. In contrast, the reverse pattern of results is observed for the students who participanted in justice-focused citizenship programmes. In these cases, students showed enhanced political interest and changed their explanations of poverty.

School practices in this domain not only vary between schools because of their idiosyncracies, but also because of the educational system curricula. A review of curricular documents of post conflict countries by Quaynor (2012), including Latin American, East Europe and African countries, reveal relevant commonalities: the role of ethnicity, lack of trust in political parties, an avoidance of controversial issues, and authoritarianism endorsement. A different curricular analysis of the Latino American Countries who participated in the International Civic and Citizenship Study (ICCS) 2009 study (Bascopé, Bonhomme, Cox, Castillo, & Miranda, 2013) shows consistency between curricular guidelines and overall country differences in social attitudes. For example, the Dominican Republic does not include any primary content guidelines on corruption practices in its curriculums, and correspondingly achieves the highest level of perceived Permissibility of Corruption among the Latin American countries within this study.

Ideological beliefs as mediators

A critical bridge between civic education and democratic attitudes is the endorsement of ideological beliefs. A multi-wave comparison study of educational attainment between adults (highest obtained degree) and diverse forms of prejudice, including sexism, racism, conducted in Germany with a representative sample of 7 waves (Zick et al., 2008) confirms that Authoritarianism and Social Dominance are mediators of effects on racism and sexism among students with different levels of attainment (Carvacho et al., 2013). These results were replicated with representative samples of other European countries, and with panel samples in Germany and Chile in the same article.

A link between lower socioeconomic status and higher authoritarianism is expected and can be seen by contrasting ideological beliefs of those with fewer vs. more years of schooling (e.g., comparing adults with no tertiary education and adults with university degrees). The contrasts between these extremes relate to the idea of "constrained political experience": while the first group is most likely underexposed to political discussion and reflection, the second is expected to have participated in such instances. There is a corresponding expectation that authoritarianism is endorsed differently according to the socioeconomic status of people. This comes from the allocation hypothesis of schooling and political socialization (Ichilov, 1991). Education serves the purpose of giving a social position to differing strata of the social hierarchy, which in turn conveys certain values during the process (Stubager, 2008, 2009).

However, just as we noted earlier with respect to the prediction of democratic attitudes, a more specific way to assess the effects of education on ideological beliefs (e.g., lower authoritarianism) is with a direct measure of the political knowledge acquired by individuals in the course of their education. The International Civic and Citizenship Education Study (ICCS) provides such a measure, which covers understanding and competencies of civic and citizenship (Schulz et al., 2008). This measure ranks students by their level of sophistication: in the lowest end of the scale, students can justify voluntary voting in the context of freedom of political expression, and identify that democratic leaders should be aware of the needs of the people over whom they have authority. In contrast, students located at the highest end of the scale can explain much more complex matters, such as being able to evaluate a policy with respect to equality, to justify the separation of powers between the judiciary and the parliament, and to understand the reasons to regulate mass media to ensure openness for press (Schulz et al., 2010). The Latin American and Asian reports from the ICCS 2009 study find a negative relations between Civic Knowledge and Authoritarianism (Fraillon et al., 2012; Schulz, Ainley, Friedman, et al., 2011).

In the ICCS 2009 study, the measures of authoritarianism for the Latino American and the Asian countries focus on the facets of submission to authority and authoritarian aggression (Duckitt et al., 2010; Funke, 2005; Mavor, Louis, & Sibley, 2010). The first component assesses "attitudes favouring uncritical, respectful, obedient support for existing societal or group authorities and institutions", while the second component concerns agreement with the use of force and coercion (Duckitt et al., 2010, p. 690). An example of the first component is the item "The most important opinion of a country should be that of the president" (Schulz, Ainley, Friedman, et al., 2011, p. 94), and examples of the second component are "It is acceptable for the government to break the law when it considers it necessary" (Fraillon et al., 2012, p. 81) and "People in government must enforce their authority even if it means violating the rights of some citizens" (Schulz, Ainley, Friedman, et al., 2011, p. 94).

Most of the research in authoritarianism is dominated by a focus on the relationship between these factors and intergroup attitudes (e.g., Asbrock et al., 2010;

Guimond, Dambrun, Michinov, & Duarte, 2003; Johnson, Terry, & Louis, 2005; Sibley & Duckitt, 2008; Zick et al., 2008). However, the original work of Adorno (Adorno, Frenkel-brunsivik, Levinson, & Sanford, 1950), the subsequent work of Altmeyer (1981), and more contemporary work (e.g., Jost, Federico, & Napier, 2009) have pointed to the broad relation of authoritarianism with several political attitudes. Thus, potential outcome variables of interest include: political orientation (Napier & Jost, 2008), support for democratic values in general (Canetti-Nisim, 2004), conformity (Feldman, 2003; Gorfein, 1961), and unethical behaviour (Salamzadeh, 2012; Son Hing, Bobocel, Zanna, & McBride, 2007; Tan, Liu, Zheng, & Huang, 2015), among others. Given the broad scope of influence of this factor, in the present study we estimate the relationship between civic knowledge and authoritarianism to a range of different democratic attitudes, including views on sexism, views on racism, support for freedom of speech, and the perceived permissibility of corruption.

The role of school climate: open classroom discussions at school

The indirect connection between civic knowledge and democratic attitudes via (lower) authoritarianism may be different depending on features of the school. In particular, Campbell (2008) and Godfrey and Grayman (2014) attribute a compensatory role to open classroom discussion. Specifically, it is expected that students from disadvantaged backgrounds in schools with greater open discussion would benefit more from civic education, in comparison with their peers in schools with less openness. In the present study, we extend this argument by estimating if/how the relationship between civic knowledge and authoritarianism is moderated by different levels of exposure to classroom discussion. Thus, our conceptual model can be expressed by the following figure[12]:

Figure 4.1: Conceptual Model

Campbell (2008) and Hibbing & Theiss-Morse (2002), asserts that students do not learn how to become good citizens just by acquiring knowledge. These authors

[12] Bear in mind that conceptual model, does not express exhaustively the model specification used in the present study. For the latter, see Figure 4.2.

attribute a crucial role to openness of discussion within the classroom (showing that teachers 'walk the talk'), in order to foster critical thinking (Godfrey & Grayman, 2014) and to embrace conflict as part of the democratic process (Campbell, 2008). Thus, we hypothesized that students with higher exposure to classroom discussion would be more critical of authority and more tolerant (Godfrey & Grayman, 2014; Hanh, 1998; Newmann, 1991), and as a consequence, would be less likely to endorse authoritarianism. This rationale is consistent with the work of Van Hiel, Pandelaere, and Duriez (2004), which suggests that educational interventions directed to reduce the 'need for closure' (i.e., a form of cognitive conservatism, and close mindedness) might reduce authoritarianism. High need for closure fosters opinion uniformity, autocratic leadership and in-group favouritism (Kruglanski, Pierro, Mannetti, & De Grada, 2006). We argue that variations in school climate factors, in terms of open classroom discussion, offer a natural experiment to test these relations in real-life settings.

The Present Study

We use secondary data from the ICCS 2009 study (Brese, Jung, Mirazchiyski, Schulz, & Zuehlke, 2011a), which collects representative student samples within schools from various countries. We estimate a model of democratic attitudes in which civic knowledge is a predictor, authoritarianism is a mediator, and open classroom discussion is a moderator. In separate models, gender equality, support for equal rights of all ethnic groups, freedom of speech support, and permissibility of corruption were specified as outcomes. All estimates were controlled for the socioeconomic background of the students.

Method

Data sources, measures, and participants

We retrieved the measures of authoritarianism[13] (AUTGOV, UNDEMGOV), and permissibility of corruption (ATTCORR, CORRPUB) respectively, from the International Civic and Citizenship Study (ICCS) 2009 study, from Latin-America (Schulz, Ainley, Friedman, et al., 2011) and Asian (Fraillon et al., 2012) regional modules. From the international module, we retrieved the measures of support for gender equality (GENEQL), support for ethnic rights (ETHRGHT), freedom of speech support (DEMVAL), civic knowledge (PV1CIV-PV5CIV), open classroom discussion (OPDSIC) and socio -economic status of the students (NISB). A detailed report on the construction of each questionnaire scale can be found in Appendix Paper 4, section F.

This study includes representative samples of six Latin American countries (Chile, Colombia, Guatemala, Mexico, Paraguay, and Dominican Republic) and five countries from the Asian region (Taiwan[14], Hong Kong SAR, Indonesia, Korea Republic, and Thailand). We selected these samples because they were the only ones which included the full array of relevant measures (for details of items, see Brese, Jung, Mirazchiyski, Schulz, & Zuehlke, 2011b, 2011c), which excluded representative samples of the European countries.

The sampling frame of (ICCS) 2009 study targets all students at grade 8, with a mean age of 13.5 years at the age of testing, and entails an effective sample size of 400 observations per country, by sampling nearly 150 schools and over 3000 students per participating country (see Schulz, Ainley, & Fraillon, 2011 for more details).

Analysis Plan

We present a model driven by theory, which offers an explanatory mechanism for the relationship between civic education and democratic attitudes. A direct effect and first stage moderated mediation model (K. J. Preacher et al., 2007) was fitted, using

[13] In parentheses, we indicate the official names of the variables, which were used as they appeared in the official data release, in order to facilitate replication. To consult individual items per scale, please check the Appendix (Paper 4), section F.

[14] The official name used in the ICCS 2009 Asian report (Fraillon et al., 2012), was Chinese Taipei. We preferred the term Taiwan in order to explicitly refer to the whole educational system of Taiwan, and avoid confusion.

Mplus (L. K. Muthén & Muthén, 2012) and Mplus Automation (Hallquist & Wiley, 2012) to extract parameters of all fitted models. We are accounting for: nested structure, complex sample design (Stapleton, 2013) and plausible values (L. Rutkowski et al., 2010). Variance estimation was conducted via Taylor series Linearization (Asparouhov & Muthén, 2010), which yields similar results to Jackknife variance estimation with large samples (Stapleton, 2008), and additionally generates Chi Square fit statistics for the estimated models.

Below, we focus our attention on the main results of two sets of parameters: a) the decomposition of the total effect of civic knowledge on each outcome; and b) the estimation of the linear moderated mediation index. The first set of parameters permits to test if civic knowledge is indirectly related to the outcomes via authoritarianism. The estimates of the linear moderated mediation index assess if this indirect effect varies systematically at different levels of open classroom discussion, the hypothesized moderator (Hayes, 2015). Thus, direct effect and indirect effect estimates permit us to address two key questions: what is the relationships between civic knowledge and different democratic values? and how and why are these factors linked? Furthermore, the linear moderated mediation index provides a way of investigating the role of open classroom discussion, given our hypothesis that it moderates the relation between civic knowledge and authoritarianism.

Because open classroom discussion is a reflective measure of the school environment (Lüdtke et al., 2008), and not a classical individual difference measure, we decomposed this factor into two components: the within cluster variation, and the between cluster variation, in the manner reported by Campbell (2008). Specifically, the frame of reference for open classroom discussion is the school as whole and not the experience of students as individuals; it rather captures the experience of students as a collective, i.e. as a group. These two components are necessarily orthogonal, because centering measures within cluster purges between variance among clusters (Enders & Tofighi, 2007). This specification permits us to inspect the effects of open classroom at the cluster level (Godfrey & Grayman, 2014), and jointly estimating if within-cluster variations on this measure makes any difference.

Additionally, we wanted to estimate if open classroom discussion exposure interacts with (i.e., moderates) effects of civic knowledge. Thus, we included interaction terms for these variables. Our model specification resembles model 10 in Hayes's templates (Hayes, 2013). These parameters enable us to estimate jointly the indirect effects of civic knowledge on the outcomes, via Authoritarianism, while addressing the conditional effect of the moderator. Furthermore, with this specification we can get a linear estimate of moderation of the indirect effect under inspection (Hayes, 2015).

We decomposed socioeconomic status of students into within and between components in the manner described above for open classroom discussion. That is, we created a variable centred at the cluster level to derive within-school variations, and created a second variable of school means centred at the grand mean to estimate effects of school mean socioeconomic background. This procedure enabled us to control the estimates of the compositional effects of schools on civic knowledge attainment between students (e.g., Collado, Lomos, & Nicaise, 2014).

A graphical representation of the estimated model is display in Figure 4.2. In this diagram, each coefficient presents a label, which allows us to calculate the estimates of the indirect effects and linear moderated mediation index following Hayes (2013, 2015) and Preacher and colleagues (K. J. Preacher et al., 2007).

Figure 4.2: Estimated model

All covariates were centred to the grand mean of each country, with the exception of the within cluster variation components of SES and open classroom discussion, as discussed above. Thus, the intercept of the outcome of each model can be interpret as the corrected mean, when all covariates are at the grand mean of each country. Additionally, we divided civic knowledge scores by 10, jointly to grand mean centering, in order to get coefficients of similar size to the rest of the covariates. Civic knowledge, is the only measure whose scale has an intended international mean of 500 and a standard deviation of 100 points, while the rest of the measures, excluding SES, have an expected mean of 50 and a standard deviation of 10 points (see Appendix (Paper 4), section F).

We used the current model specification, because a regression approach, even though it does allow us to assess partially the first and second question of our research, is not sufficient to assess our third question: to estimate the indirect effect at conditional levels of the moderator. However, we include population estimates of the regression approach in Appendix (Paper 4), section G. A second alternative is to specify a path analysis, similar to the conceptual model depicted in Figure 4.1. Yet, we know this model does not address the issue of the between and within components of the compositional effect of SES on civic knowledge, nor the reflective nature of the open classroom discussion. Additionally, we have estimated such a specification previously and we know this does not fit the data well. Finally, we have chosen a fixed effect model specification (i.e. design based method), instead of a model based method (i.e., multilevel model, mixed model, random effects). This option was chosen because it was a more parsimonious option: these models produce fewer parameters. These two family of models can produce almost equivalent estimates under certain conditions such as large sample size per cluster, and large number of clusters (Begg & Parides, 2003; McNeish, 2014). The chosen model expresses the relative contribution of each factor from deviations of the population grand mean of each outcome (i.e., the intercepts), with no references to the differences between clusters, as it would be in multilevel models. Other model specifications are discussed in the discussion section.

In the present study, we compared different outcomes, different countries, from different regions with different measures of the mediator and one of the outcomes, with the aim of assessing how robust our proposed model is. Given that our final research questions concerns the interaction of measures (civic knowledge and open classroom discussion) between schools, subdividing the sample within country into groups is a less convenient option because of the loss of power. In contrast, using different representative samples of 6 countries with a common language, and 5 countries with different languages, provides a more thorough assessment the proposed model's robustness.

Results

Fit Statistics

Considering the most demanding fit statistic, the Chi Square, the results of this index indicates the model fits (χ^2 (12) <=21.03, p > .05) for all the outcome variables, among the Latin American samples, with the exception of Colombia. In contrast, the model fits less well among the Asian Region samples; in this region, it fits well across all variables only for Indonesia. Additionally, it fits well on the outcome of Ethnic

Rights for Taiwan, and it fits well for Permissibility of Corruption and Freedom of Speech Support for Thailand.

The Chi Square fit index is considered too sensitive to sample size (T. A. Brown, 2006), and less suited for large scale assessments (L. Rutkowski & Svetina, 2013). If we assess the model fit by alternative fit indices, results seem acceptable in most cases. For example, using the criteria of joint fit for CFI, RMSEA and SRMR, on the recommended thresholds for CFI =>.95, RMSEA <=.05 and SRMR <=.05 (see Schermelleh-Engel, Moosbrugger, & Müller, 2003), the model fits well for all Latin American Samples, now including Colombia, as well as for all Asian samples, with the exception of the students from Hong Kong SAR.

If we compared the models in terms of accounted variance in the outcome variables (R^2), models are more potent in explaining the domains of Gender Equality (R^2 min=.14, max=.35, average=.26), and Permissibility of Corruption (R^2 min=.24, max=.50, average=.40). In contrast, predicted variance in support for Ethnic Rights (R^2 min=.03, max=.17, average=.11) and Freedom of Speech Support (R^2 min=.07, max=.22, average=.11) is lower. Also, similar to the previous indices, the models account for more variance among the Latin American samples than among the Asian Region samples.

Table 4.1: Fit Statistics Summary

Decomposition of the Civic Knowledge total effect

In all cases, students with more civic knowledge attainment endorsed more gender equality, ethnic rights, and freedom of speech support and were less permissive of corruption. As well as any direct effects, all tested samples presented indirect effects via authoritarianism that were significantly different from zero (see Table 4.2). In other words, the total effect of civic knowledge on each outcome occurred at least partially due to its negative relation with authoritarianism. The specific effects we are estimating account for varying levels of open classroom discussion, which can vary between and within schools. In particular, the reported indirect effects of civic knowledge are the estimates when the (mean-centred) moderators are zero, and hence represent the indirect effects for students with average levels of open classroom discussion per country in each model (see equation formulas in Appendix (Paper 4), sections A and B).

Table 4.2: Civic Knowledge Total Effect decomposition for each outcome

----

If direct, indirect and total effects are averaged across models, one can estimate a rough percentage of variation attributable to indirect effects by dividing the indirect effect by the total effect. This rough estimate[15] indicates that indirect effects of civic knowledge are largest for permissibility of corruption (51.3%), followed by gender equality (25.3%), with freedom of speech support and ethnic rights exhibiting smaller percentage of variations explained in this way (8.3% and 7.9%, respectively). Full model estimated parameters are provided in Appendix (Paper 4), section E, per each country and each outcome.

It should be noted that some inconsistent parameters were observed for the prediction of ethnic rights and freedom of speech support models in Indonesia, Thailand and Dominican Republic. These indirect effects are negative, which is opposite to the direction of the expected effect. We inspected the matrix of correlations by country and descriptives. We found that these countries presented the smallest correlations between

[15] A more accurate version of this estimate across studies would weight the vector of parameters by the reverse proportional of its variance, to resemble a meta-analytic estimate. However, the main substantive conclusion highlighted in this text would remain the same.

authoritarianism and these outcomes (see Figure 4.3 and Appendix (Paper 4), section G) and the highest country level means on authoritarianism. This may suggest possible ceiling effects on the mediator.

Moderated Mediation

To assess if the indirect effects vary at different levels of the moderator, we estimated the Linear Moderated Mediation Index (Hayes, 2015). This index helps us to assess if indirect effects vary as a function of the moderator, and hence, if the indirect effects are larger or smaller under different values of the moderator. However, because the scale of this parameter is on the scale of the slope coefficients it can be difficult to interpret at face value. Therefore, we created a table in which we coupled the indirect effects of civic knowledge for each outcome, at high and low levels of open classroom discussion, at +/- 10 points from the grand mean, which is the internationally expected standard deviation (SD) for this scale (Schulz, Ainley, & Fraillon, 2011). These parameters express how the size of the indirect effect varies at different levels of the moderator, at the school level. The linear moderated mediation index (LMM) is the slope of this relation.

We defined the linear moderated mediation index as the multiplication of the coefficient of the interaction between school open classroom discussion and civic knowledge in predicting the outcome variable, hence the parameters a5 and b1 from the diagram in Figure 4.2. Note that these parameters are purged of the within-cluster variation of students at the classroom level of the moderator, and thus can be interpreted as the linear relation of the different indirect effects at varying values of the moderator at its specific inference level (i.e., the school). Similarly, we estimated the indirect effects at high “(a1+a5*10)*b1” and low “(a1+a5*-10)*b1” levels of the moderator, thus giving us indirect effects of civic knowledge on democratic attitudes when classroom discussion is high vs. low (at the within- and between-school levels).

Table 4.3 Indirect Effects at High and Low levels of open classroom discussion

----

Results indicate that the linear moderated mediation index is different from zero for 7 out of the 11 countries, on all included outcomes. These includes all Latin

American samples, with the exception of Chile, and additionally the samples of Indonesia and Thailand.

If there is something in common between these countries, it is their differing mean country levels of civic knowledge and authoritarianism. This can be shown with a scatter plot (see Figure 4.2). In this figure, we plotted the mean country score of authoritarianism on the y-axis, and on the x-axis, we plotted the mean country score of civic knowledge. Interestingly, all the countries which presented evidence in favour of moderation of the indirect effect, were below the expected international average of Civic Knowledge (500 pts), and/or were above the international expected average Authoritarianism (50 pts) (Schulz, Ainley, & Fraillon, 2011).

Figure 4.2: Country-level means in civic knowledge and authoritarianism,

marked by evidence of moderated mediation

Discussion

We believe the proposed models provide evidence in line with the learning hypothesis (Caro & Schulz, 2012) regarding the power of civic education to explain democratic attitudes. Furthermore, they show that the link between civic knowledge and attitudes is partially mediated by ideological beliefs. And finally, they provide evidence that these indirect effects vary under differing levels of exposure to open discussion within the classrooms.

Broadly speaking, our models imply that differing levels of political attitudes are related to political sophistication (i.e., civic knowlege), and are partially mediated by students' ideological beliefs (i.e., endorsement of authoritarianism), after controlling for variations in socioeconomic conditions. In the current model specification, the effect of socioeconomic conditions ocurs only via civic knowledge, as this effect is fully mediated by civic knowledge (see Apendix (Paper 4), section G, SES factors are close to zero, when Civic Knowledge is included in the regression models). Moreover, schools' level of open classroom discussion appeared to moderate the indirect effect of civic knowledge on democratic attitudes, in many though not all samples. We can speculate that this occurs via the mechanism of appreciation of conflict, as suggested by Campbell (2008) and Van Hiel, Pandelaere, and Duriez (2004). Specifically, school interventions which tackle the need for closure, a sort of epistemological conservatism, may serve to accentuate the impact of civic knowledge on reducing authoritarianism, and thereby indirectly improve students' social attitudes. These results are in line with this interpretation, in the sense that authoritarianism is an ideological belief which may vary in its relations depending on social-contextual effects (Guimond et al., 2003).

The current study has a number of limitations. Each of the outcomes studied here has its own idiosyncrasies and elaborate theoretical frameworks, which we have not addressed for the sake of parsimony. For example, the contact hypothesis and social identity theory (R. Brown, 2010) have prominent roles in the prejudice literature. Similarly, social role theory (Diekman & Schneider, 2010) has much to add to the research on gender parity. However each of the estimated models is theoretically consistent with the authoritarianism framework, with respect to the role of authoritarianism in generalized prejudice (Carvacho et al., 2013; Zick et al., 2008, 2011), support for democratic rights such as freedom of speech (Cohrs, Kielmann, Maes, & Moschner, 2005), and unethical behaviour (Son Hing et al., 2007).

Authoritarianism is thought to be caused by a politically constrained experience, due to low educational attainment and associated cultural and intellectual restrictions (Carvacho et al., 2013). This leaves students from disadvantaged backgrounds at higher risk of developing this worldview and consequently endorsing generalized prejudice. However, in all samples, students with higher civic knowledge displayed higher endorsement of gender equality and ethnic equality, partially via less endorsement of authoritarianism. Further research can aim to assess how this model holds at different values of socioeconomic status to assess possible interactions with this factor.

Cohrs et al. (2005) shows that feelings of threat interact with authoritarianism, which leads individuals to endorse the limitation of civil liberties. People with higher levels of authoritarianism are more submissive to the political order, and more sensitive to any threat to the status quo (Jost et al., 2009). In contrast, students with higher levels of civic knowledge were consistently stronger supporters of freedom of speech, partially via lower authoritarianism.

Son Hing et al.'s (2007) demonstrate that high right-wing authoritarian individuals tend to comply to unethical behaviour of dominant individuals, hence they become more lenient and permissive to unethical behaviour. Thus, via conformity and submission, high endorsement of authoritarianism can be linked to permissiveness of corruption. Our results are in line with this interpretation: students with higher civic knowledge were less permissive of corruption, and almost half of the total effect occurred because of the indirect effects of civic knowledge via authoritarianism.

Crucially, in 7 of 11 countries, these indirect effects were higher when schools presented higher levels of open classroom discussion of political issues. This evidence of moderated mediation points to the significance of school climate as a key socio-contextual factor that modulates the associations between student knowledge and beliefs on the one hand, and democratic attitudes on the other. The fact that the linear moderated mediation is present mainly across countries with higher authoritarianism is consistent with previous literature (Asbrock, Christ, Duckitt, & Sibley, 2012), in which it has been found that contact interventions, for example, have a greater effect on students with higher levels of authoritarianism (and smaller effects for students with less authoritarian values). However, the lack of strict equivalency between the measures of authoritarianism and permissibility of corruption between the countries of each region adds noise to the estimates of the effects, confounding the overall estimate for all samples between regions. We should also note that the results are correlational in nature, hence conclusive inferences about causal influence are not guaranteed. Yet, the moderated indirect effects could be estimated within a causally defined framework, also called potential outcome framework (Caro, 2014; Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010). Future research could point in this direction, and estimate the expected losses and gains from attending schools with high vs. low levels of open classroom discussion, while addressing the possible effects of endogeneity between the mediator and the outcome within the model.

The specification of a fixed effects model was chosen in the present study because it is a more parsimonious option among other model specification alternatives (e.g., multilevel modes). As was mentioned earlier, these models can reach almost equivalent estimates under certain conditions such as large sample size per cluster, and large number of clusters (Begg & Parides, 2003; McNeish, 2014). However, because of the small size of the coefficient of the linear moderated mediation, in further research we aim to assess this conceptual model under a more robust specification, such as multilevel latent covariate model, which enables us to handle the interaction between two level 1 measures, to assess an interaction at both levels (Ryu, 2015).

One the main limitations of a linear moderated mediation is the interpretation of its estimates. This index is in the scale of slope of the estimated parameters. However, that given all the studied outcomes come from an item response theory scaling, the estimated effects can be plotted against the wright map of the measurement model (Torres Irribarra & Freund, 2014). It is possible to connect the estimated parameters with the expected values of the studied outcome, using the modelled pattern of responses of the population of students. Thus, the current estimated model can be linked with its measurement model to facilitate interpretation.

Simpler model specifications might be easier to communicate, yet these convey the risk of missing relevant effects of substantive theory. In the particular case of education, such practice might involve losing some plausible recommendations for public policy and for teaching practice. Different model specifications lead to different conclusions, and so when relevant factors and effects are ignored, no compelling recommendation regarding this dimension can be made. This is especially the case with school climate factors, which under different specification might not show their true value. The present study proposes a feasible way to address this issue, by specifying open classroom discussion as within and between components, which helped us to expose the relations between civic knowledge and authoritarianism endorsement using a moderated mediation approach.

This research is relevant because it sheds light on the relationship of civic knowledge and civic education, gives a partial explanation of how these are linked to the endorsement of relevant democratic values, such as gender equality, support for ethnic rights, support of freedom of speech and lower permissibility of corruption. Finally, our results offer recommendations for teaching practice relating democratic values. Specifically, the discussion of political issues within the classroom may encourage students' reflection on these topics and foster different points of view, thus laying a foundation for endorsing democratic attitudes.

We believe the present results have much in common with other research programs, such as implicit theories of learning (Dweck, Chiu, & Hong, 2009), and interpersonal expectancy beliefs (Rosenthal, 2002). All of these highlight the central role of beliefs in the study of educational outcomes. This study is an example of how "worldviews" are fundamental in democratic attitudes, and how the school's social process matters in this regard. Our results set an agenda for further research to extend our analyses and evaluate in more depth these patterns of socio-contextual effects.

Figures and Tables Paper 4

Figure 4.1 Conceptual Model

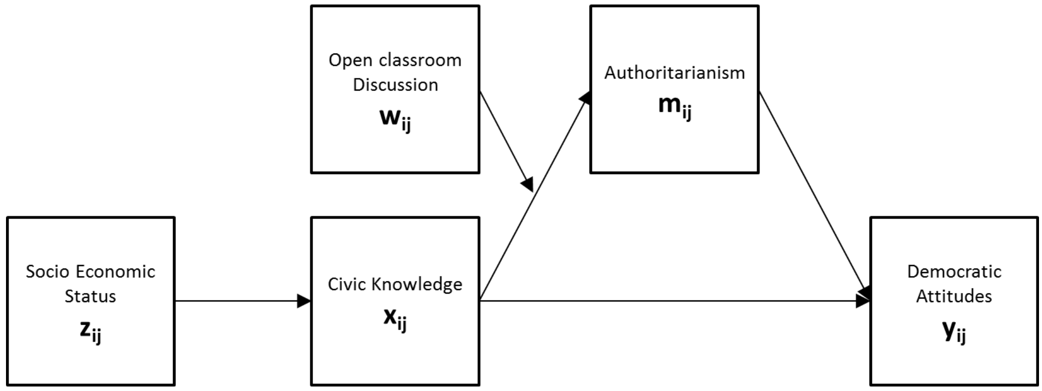
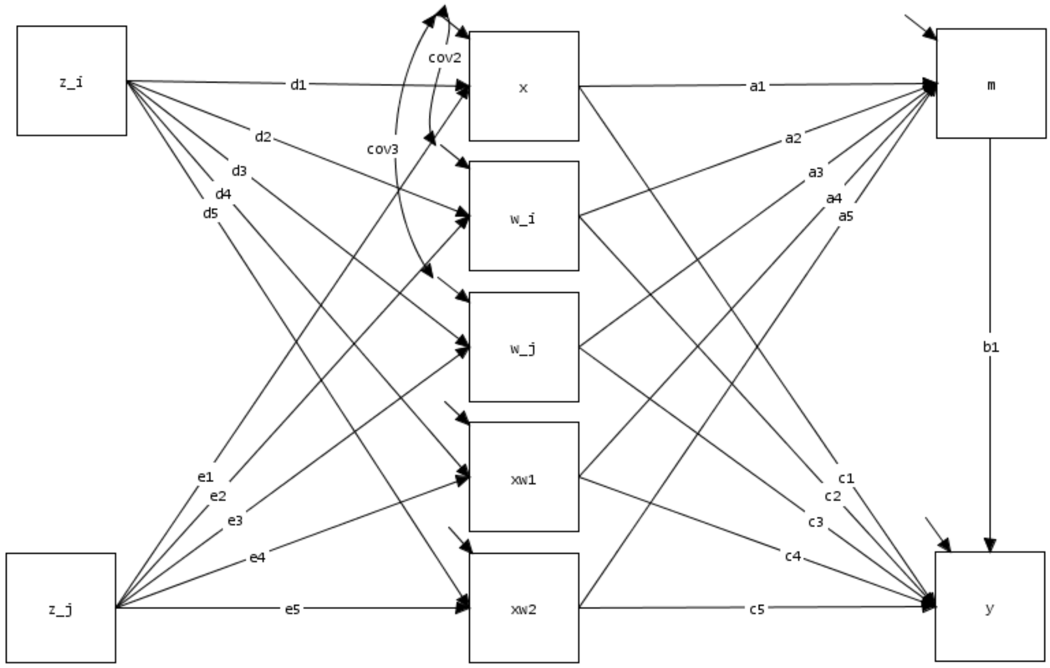


Figure 4.2 Estimated model



Notes:

z_i = within cluster deviation from the cluster mean on socio economic status (NISB).
z_j = between cluster deviations from the grand mean on socioeconomic status (NISB). School level means of socio economic status, centered to the grand mean.
x = students' Civic Knowledge (PV1CIV, PV2CIV, PV3CIV, PV4CIV, PV5CIV) centred at the grand mean
w_i = within cluster deviation from the cluster mean on Open Classroom Discussion (OPDISC). Student's deviation of Open Classroom Discussion from the cluster mean.
w_j = between cluster deviations from the grand mean on Open Classroom Discussion (OPDISC). School level means of Open Classroom Discussion, centered to the grand mean.
xw1= interaction term of civic knowledge and Student's deviation of Open Classroom Discussion from the cluster mean (x*w_i).
xw2 = interaction term of civic knowledge and School's means of Open Classroom Discussion (x*w_j).
m = Authoritarianism, centered at the grand mean (AUTGOV in Latino Samples, and UNDEMGOV in Asian Samples).
y = corresponding measures outcome on each model:
Gender Equality Support (GENEQL)
Ethnic Equality Support (ETHRIGT)
Permissibility of Corruption (ATTCORR in Latino Samples, and CORRPUB in Asian)
Freedom of Speech Support (DEMVAL)

Table 4.1 Fit Statistics Summary

	Gender Equality	Ethnic Rights	Permissibility of Corruption	Freedom of Speech
Absolute Fit				
Asian Samples				
Hong Kong SAR	No	No	No	No
Indonesia	Yes	Yes	Yes	Yes
Republic of Korea	No	No	No	No
Thailand	No	No	Yes	Yes
Taiwan	No	Yes	No	No
Latino Samples				
Chile	Yes	Yes	Yes	Yes
Colombia	No	No	No	No
Dominican Republic	Yes	Yes	Yes	Yes
Guatemala	Yes	Yes	Yes	Yes
Mexico	Yes	Yes	Yes	Yes
Paraguay	Yes	Yes	Yes	Yes
Alternative Fit				
Asian Samples				
Hong Kong SAR	No	No	No	No
Indonesia	Yes	Yes	Yes	Yes
Republic of Korea	Yes	Yes	Yes	Yes
Thailand	Yes	Yes	Yes	Yes
Taiwan	Yes	Yes	Yes	Yes
Latino Samples				
Chile	Yes	Yes	Yes	Yes
Colombia	Yes	Yes	Yes	Yes
Dominican Republic	Yes	Yes	Yes	Yes
Guatemala	Yes	Yes	Yes	Yes
Mexico	Yes	Yes	Yes	Yes
Paraguay	Yes	Yes	Yes	Yes
Accounted Variance				
Asian Samples				
Hong Kong SAR	0.14	0.10	0.24	0.08
Indonesia	0.16	0.12	0.29	0.13
Republic of Korea	0.15	0.10	0.30	0.22
Thailand	0.27	0.12	0.36	0.12
Taiwan	0.27	0.11	0.33	0.13
Latino Samples				
Chile	0.31	0.13	0.45	0.08
Colombia	0.35	0.15	0.43	0.10
Dominican Republic	0.27	0.03	0.50	0.08
Guatemala	0.34	0.13	0.49	0.09
Mexico	0.31	0.17	0.50	0.13
Paraguay	0.34	0.09	0.49	0.09

Notes. Absolute Fit: if $\chi^2(12)<21.026$, $p<.05$ then Yes, otherwise No; Alternative Fit: if CFI>=.97 & RMSEA <=.05 & SRMR <= .05, then Yes, otherwise No. Accounted variance is estimated R^2 for each corresponding outcome, from the models.

Table 4.2 Civic Knowledge Total Effect decomposition for each outcome

		Gender Equality		Ethnic Rights		Permissibility of Corruption		Freedom of Speech	
Asian Samples		E	SE	E	SE	E	SE	E	SE
Hong Kong SAR	Total	.28	-.03 **	.23	-.03 **	-.30	-.02 **	.24	-.02 **
	Direct	.23	-.03 **	.21	-.03 **	-.21	-.02 **	.21	-.02 **
	Indirect	.04	-.01 **	.02	-.01 **	-.09	-.01 **	.03	-.01 **
Indonesia	Total	.29	-.02 **	.23	-.02 **	-.46	-.03 **	.34	-.03 **
	Direct	.27	-.02 **	.29	-.03 **	-.33	-.02 **	.39	-.03 **
	Indirect	.02	-.01 **	-.06	-.01 **	-.13	-.01 **	-.05	-.01 **
Republic of Korea	Total	.36	-.02 **	.35	-.02 **	-.34	-.02 **	.59	-.02 **
	Direct	.29	-.02 **	.29	-.02 **	-.19	-.02 **	.52	-.02 **
	Indirect	.07	-.01 **	.06	-.01 **	-.16	-.01 **	.07	-.01 **
Thailand	Total	.38	-.02 **	.23	-.02 **	-.49	-.02 **	.29	-.03 **
	Direct	.34	-.02 **	.30	-.02 **	-.30	-.02 **	.35	-.03 **
	Indirect	.05	-.01 **	-.07	-.01 **	-.19	-.01 **	-.06	-.01 **
Taiwan	Total	.44	-.02 **	.26	-.02 **	-.44	-.02 **	.36	-.02 **
	Direct	.34	-.02 **	.22	-.02 **	-.26	-.01 **	.32	-.02 **
	Indirect	.10	-.01 **	.04	-.01 **	-.17	-.01 **	.04	-.01 **
Latin Samples									
Chile	Total	.49	-.02 **	.29	-.02 **	-.60	-.02 **	.23	-.02 **
	Direct	.34	-.03 **	.23	-.03 **	-.24	-.02 **	.16	-.03 **
	Indirect	.15	-.01 **	.06	-.02 **	-.36	-.01 **	.07	-.01 **
Colombia	Total	.51	-.02 **	.29	-.02 **	-.53	-.02 **	.28	-.03 **
	Direct	.32	-.02 **	.22	-.02 **	-.22	-.02 **	.20	-.04 **
	Indirect	.19	-.01 **	.06	-.01 **	-.31	-.01 **	.08	-.01 **
Dominican Republic	Total	.46	-.03 **	.15	-.04 **	-.55	-.03 **	.34	-.04 **
	Direct	.35	-.03 **	.20	-.04 **	-.19	-.03 **	.38	-.05 **
	Indirect	.11	-.01 **	-.05	-.02 **	-.36	-.02 **	-.04	-.01 **
Guatemala	Total	.54	-.03 **	.32	-.04 **	-.68	-.03 **	.28	-.02 **
	Direct	.37	-.03 **	.26	-.04 **	-.37	-.03 **	.24	-.03 **
	Indirect	.18	-.01 **	.05	-.01 **	-.31	-.02 **	.04	-.02 *
Mexico	Total	.38	-.01 **	.35	-.02 **	-.64	-.02 **	.37	-.02 **
	Direct	.26	-.01 **	.29	-.02 **	-.22	-.02 **	.29	-.02 **
	Indirect	.12	-.01 **	.06	-.01 **	-.43	-.02 **	.07	-.01 **
Paraguay	Total	.47	-.03 **	.19	-.03 **	-.51	-.03 **	.29	-.03 **
	Direct	.33	-.03 **	.14	-.03 **	-.18	-.03 **	.23	-.03 **
	Indirect	.14	-.01 **	.05	-.01 **	-.34	-.02 **	.06	-.01 **

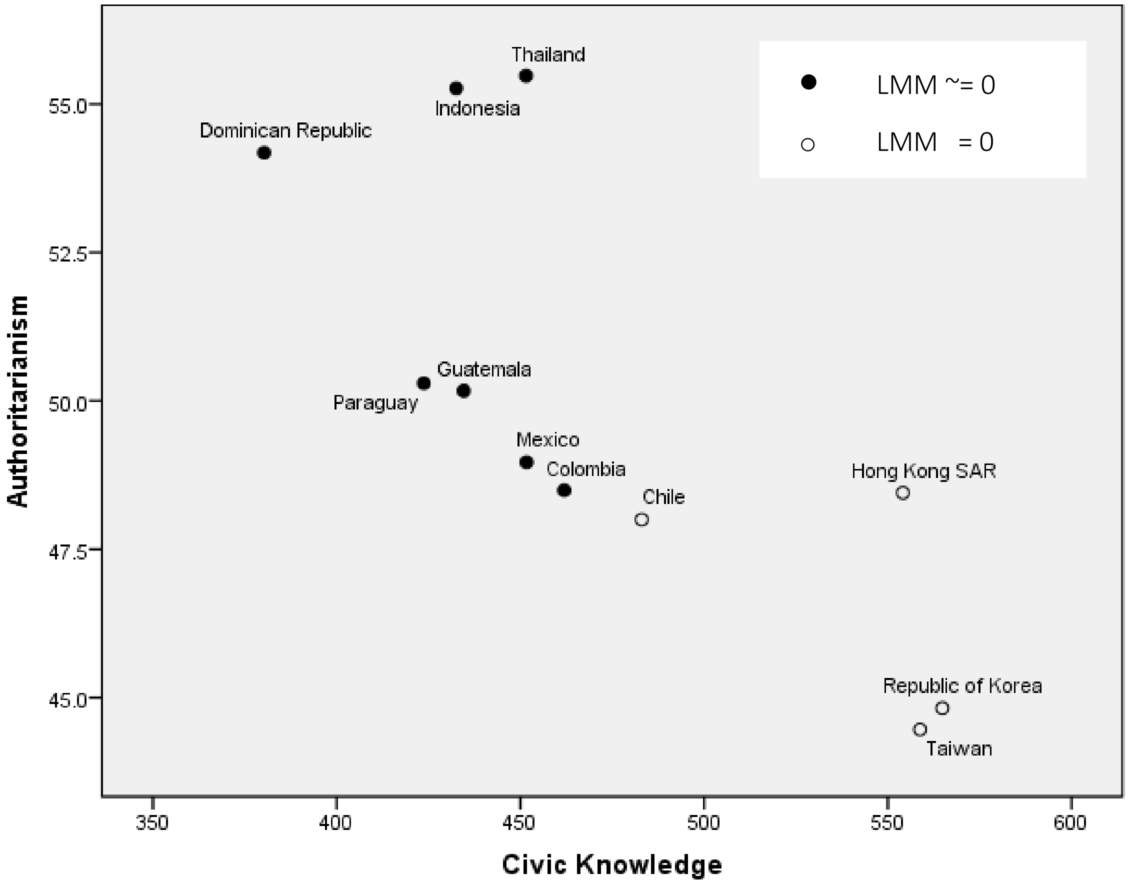
Note: E = estimate, SE=standard Errors, ** p<.01, * p<.05. Unstandardized estimates

Table 4.3 Indirect Effects at High and Low values of open classroom discussion

		Gender Equality		Ethnic Rights		Permissibility of Corruption		Freedom of Speech	
Asian Samples		E	SE	E	SE	E	SE	E	SE
	High	.06	-.02 **	.03	-.01 *	-.14	-.03 **	.05	-.02 **
Hong Kong SAR	Low	.02	-.01 *	.01	-.01 *	-.05	-.03	.02	-.01 *
	LMM	.00	.00	.00	.00	.00	.00	.00	.00
Indonesia	High	.03	-.01 **	-.10	-.02 **	-.20	-.03 **	-.08	-.02 **
	Low	.01	-.01	-.03	-.01 *	-.06	-.03 *	-.02	-.01 *
	LMM	.00	.00 *	.00	.00 **	-.01	.00 **	.00	.00 *
	High	.07	-.01 **	.06	-.02 **	-.16	-.03 **	.08	-.02 **
Republic of Korea	Low	.07	-.01 **	.06	-.01 **	-.15	-.03 **	.07	-.01 **
	LMM	.00	.00	.00	.00	.00	.00	.00	.00
Thailand	High	.07	-.01 **	-.11	-.02 **	-.31	-.03 **	-.10	-.02 **
	Low	.02	-.01 **	-.03	-.01 **	-.08	-.03 **	-.03	-.01 **
	LMM	.00	.00 **	.00	.00 **	-.01	.00 **	.00	.00 **
Taiwan	High	.10	-.01 **	.04	-.01 **	-.18	-.02 **	.04	-.01 **
	Low	.10	-.01 **	.04	-.01 **	-.17	-.03 **	.04	-.01 **
	LMM	.00	.00	.00	.00	.00	.00	.00	.00
Latin Samples									
Chile	High	.15	-.01 **	.06	-.02 **	-.36	-.02 **	.07	-.01 **
	Low	.15	-.01 **	.06	-.02 **	-.36	-.03 **	.07	-.01 **
	LMM	.00	.00	.00	.00	.00	.00	.00	.00
Colombia	High	.23	-.02 **	.08	-.02 **	-.39	-.03 **	.10	-.02 **
	Low	.14	-.02 **	.05	-.01 **	-.23	-.03 **	.06	-.01 **
	LMM	.01	.00 **	.00	.00 *	-.01	.00 **	.00	.00 *
	High	.17	-.03 **	-.08	-.02 **	-.55	-.06 **	-.06	-.02 **
Dominican Republic	Low	.05	-.02 *	-.02	-.01	-.16	-.06 *	-.02	-.01
	LMM	.01	.00 **	.00	.00 *	-.02	-.01 **	.00	.00 *
Guatemala	High	.24	-.03 **	.07	-.02 **	-.43	-.04 **	.05	-.02 *
	Low	.11	-.02 **	.03	-.01 **	-.20	-.04 **	.02	-.01 *
	LMM	.01	.00 **	.00	.00 *	-.01	.00 **	.00	.00 *
Mexico	High	.15	-.02 **	.08	-.02 **	-.54	-.05 **	.09	-.02 **
	Low	.09	-.02 **	.04	-.01 **	-.31	-.05 **	.05	-.01 **
	LMM	.00	.00 *	.00	.00 *	-.01	-.01 *	.00	.00 *
Paraguay	High	.20	-.02 **	.08	-.02 **	-.47	-.05 **	.08	-.02 **
	Low	.09	-.02 **	.03	-.01 **	-.20	-.04 **	.03	-.01 **
	LMM	.01	.00 **	.00	.00 *	-.01	.00 **	.00	.00 **

Note: High = indirect effect when Open Classroom Discussion is +1 SD over the grand mean, and Low = indirect effect when Open Classroom Discussion is -1 SD below the grand mean, LMM = index of linear moderated mediation, E = estimate, SE=standard Errors, ** p<.01, * p<.05. Unstandardized estimates.

Figure 4.3 Country-level means in civic knowledge and authoritarianism, marked by evidence of moderated mediation



Note: LMM ~= 0, stands for countries in which the linear moderated mediation was different from zero; LMM = 0 stands for countries in which the moderation of the indirect effect was not different from zero. This is the case for all studied outcomes.

General discussion

Summary of key findings

School climate as an umbrella concept. It was noted earlier that school climate is an elusive term. However, considering school climate as a multi-faceted umbrella concept, rather than a single unidimensional construct, was useful for the development of the current work. By reviewing the historical roots of the concept within educational research, and its particular role in the school effectiveness literature, it was found that different factors from the school or the classroom environment have been thought to capture Perry's (1908) initial idea of the school *atmosphere* – that intangible feature that makes school environments different from one another. Nevertheless, the myriad of school climate factors, dimensions and domains is often overshadowed by an initial focus on learning outcomes. It seems that, by a passive importation of the conceptual approaches that dominate work on school effectiveness, school climate research has often had little to say on other non-scholastic, yet relevant outcomes, such as teacher turnover, teacher job satisfaction, students' social attitudes and values.

Taking a step towards understanding the broader role played by aspects of school climate, the present work analysed two different outcomes for teachers and two different outcomes for students. In the first two papers we addressed the problem of teacher turnover in the context of understanding school differences. In paper 3, we addressed the relation between bullying and math achievement. In paper 4, we addressed the association between civic knowledge and democratic values endorsement. All of these relations occurred in clustered settings, given that teachers and students are nested within schools; because they are not randomly assigned to schools, they are subject to clustered differences. Moreover, students and teachers form clustered collectives that can potentially develop clustered variance. Thus, any relevant school outcome under investigation is subject to clustered effects phenomena. The present work suggests that consideration of school climate factors offers a valuable explanatory framework for understanding a variety of such clustered outcomes. The following section summarizes the main results regarding school climate factors and school differences, with analyses mainly focused on the Chilean schooling context (which is one marked by high levels of socioeconomic segregation).

Teacher actual turnover. In paper 1, to appropriately present our case, we first estimated population level turnover rates, under a survival model. This model framework allowed us to depict the most accurate state of affairs regarding the outcome under investigation, and identify what factors from the school context are related to it. With rigorous statistical controls, it was shown that schools serving more socioeconomically disadvantaged populations of students present higher rates of turnover among teachers in their first contract. This result was independent of teacher characteristics. Moreover, given the additive property of the estimates of this model, we could estimate the profile of teachers at greater risk of turnover: male, younger than 30 years, at the start of their career, teaching in secondary grades, in subsidized schools. In study 2, we further inquired what other factors can account for these relations. Organizational aspects of the school predicted teacher turnover, above and beyond schools' socioeconomic intake and academic achievement level. At the population level, teachers' subjective experience of principal leadership and school priorities moderated likelihood of turnover: teachers who worked in schools with a shared vision, fair treatment, and greater support were less likely to move from their schools, and teachers who experienced school managements where the relations between staff are prioritised were less likely to turnover from one year to the next. These results were moderated by teachers' experience. Turnover rates among novice teachers (0-3 years of experience) were particularly responsive to these school climate factors. School leadership and management relations were found to be protective factors against teacher turnover, whereas academic monitoring was found to be positively related to teacher turnover. In contrast, non-novice teachers' turnover rates were responsive to positive school relations among the school members. Higher rates on this dimension were negatively related to teacher turnover.

A relevant highlight of these results is that academic monitoring, a school climate factor frequently associated with greater school effectiveness (Scheerens et al., 2003), was related to *increased* teacher turnover among novice teachers. Therefore, a single school climate factor can show positive and negative relations to difference school outcomes. In the context of teacher turnover, we interpreted this effect within the job and demands resources framework (Bakker & Demerouti, 2007) as possible stressor factor.

Teacher turnover intentions and school climate factors. One of the limitations of paper 1 was the inability to capture school climate factors as clustered level effects. This was due to the fact that the sample in study 2 of this paper presented high levels of sparsity, where a high proportion of schools provided too few teachers. Thus in paper 2, we relied on the use of a large scale assessment study which sampled random schools and random teachers at the country level, namely the Teaching and Learning International Survey from 2013 (TALIS 2013). In this study we addressed the relationship between school climate factors and teacher turnover intentions. Two school climate factors were inspected, namely teacher-student relationships and classroom discipline. Turnover intentions were also modeled in a twofold manner: intentions to leave the profession and intentions to quit the school. School classroom discipline presented contextual effects. Schools with higher than average classroom discipline presented lower teacher intentions to leave the profession, regardless of individual differences in teachers' levels of satisfaction with the current work environment. Teachers' job satisfaction with the current school, on the other hand, revealed a complex pattern of associations among variables. At the within-school level, the link between job satisfaction and intentions to quit the school was negative: more satisfied teachers did not want to quit. At the aggregated level, this was also true. Yet, when the contextual effect was estimated, this effect shifted its direction. This is a paradoxical result suggesting that, above and beyond the effects for individual teachers, being in a school where job satisfaction is generally high can actually predict *higher* levels of intentions to quit the school. This a similar effect, in terms of shape, to the big fish little pond effect (Blalock, 1984; Burstein, 1980; Marsh, Abduljabbar, et al., 2014), which is also a paradoxical effect. These two results highlight the advantage of separating within and between cluster relations, as these two angles can show different results, which cannot be reduced to one another.

Crucially, based on the job and demands resources model (Bakker & Demerouti, 2007; Bakker et al., 2007) and the literature on job turnover (Berry et al., 2012; Podsakoff et al., 2007), we specified job satisfaction as a mediator of the school climate factor effects. At the aggregated level, schools with higher than average teacher-student relations showed greater teacher job satisfaction, which in turn was negatively related to intentions to leave the profession. However, most of the indirect effects were present at the within level. Job satisfaction, indeed, was the relevant mediator to explain relative differences between teachers' classroom experience and both kinds of teacher intentions of turnover, in the expected direction. Specifically, more positive teacher-student relations and classroom discipline were related to greater job satisfaction, and more satisfied teachers did not want to leave the profession or the school.

Bullying and student math achievement. In paper 3, we estimated the contextual effects of bullying rates and school mean levels of belonging and engagement, this time as potential explanations for variance in students' math achievement. These two different factors, which were measured as formative indicators by aggregating the students' individual experience, were consequently split into their within and between cluster components. Traditional multilevel estimates showed that these two factors presented contextual effects. Students in schools with higher than average levels of bullying were expected to have poorer achievement results than students with similar characteristics in other schools. Similarly, students in schools with higher levels of belonging were expected to achieve higher results in math than similar students in other schools. Yet, when school differences in overall levels of perceived safety and academic emphasis were controlled for, these contextual effects diminished below sampling error. Thus, the contextual effects of bullying seemed to be explained by observed differences between the schools on other, broader attributes. Given that these factors were formative in nature, we followed a second line of inquiry to shed light on the relationship between bullying and academic achievement. At the within-school level, students who suffered bullying were less engaged with school and less engaged with the lessons, and thereby achieved lesser results. This indirect effect was moderated by students' levels of belonging, such that the indirect effect of bullying on achievement via classroom engagement was smaller for students with higher than average levels of school belonging.

In this paper, we show how formative indicators present valid interpretations at both levels of inquiry, for clustered effects and for comparisons of relative differences at the within-school level. Moreover, these factors can display complex relations such as contextual effects, indirect effects and moderated mediations.

Students' endorsement of democratic attitudes and civic knowledge. Compared to the preceding papers, this paper presented two new features: we used more than one country to fit the specified model, and we used a hybrid model to estimate conditional indirect effects at the clustered level. What is the role of open classroom discussion in democratic attitudes endorsement? Before addressing this question, we established a model to explain the relationship between civic knowledge and democratic attitudes (greater support for gender equality, ethnic rights, and freedom of speech, and lower perceived permissibility of corruption). Following the literature on authoritarianism, and the sophistication hypothesis, we expected that civic knowledge would be negatively related to authoritarianism, and in turn that lower authoritarianism would be related to higher levels of all democratic attitudes. As such, we estimated indirect effects from civic knowledge to democratic attitudes to all outcomes, via authoritarianism. All models presented indirect effects in the expected direction, with the exception of two countries[16]. Moreover, to study the role played by having an open classroom climate, we split this covariate into its within and between parts. This is a reflective indicator of school climate. In this study, we were interested in its aggregated effect at the school level. We used this component as a modifier of the relationship between civic knowledge, authoritarianism, and democratic attitudes. This model built upon previous literature, which suggests that school interventions that aim to reduce cognitive conservatism might thereby serve to reduce authoritarianism. Additionally, open classroom discussion could play a compensatory role for disadvantaged students, so we speculated that open classroom discussion could interact with civic knowledge. Thus, schools with higher classroom discussion were expected to show a higher negative relation between civic knowledge and authoritarianism. Moreover, we expected this effect would also be reflected in the indirect effects on the studied outcomes. Our estimates supported our hypothesis for 7 out 11 countries, notably the countries with lower mean levels of civic knowledge levels and/or higher authoritarianism levels.

In this paper, a single school climate factor was specified as the modifier of a relationship (a moderator), within an indirect effect chain. We used a theory driven model to create this model specification, and to test this model onto 11 countries and onto 4 different outcomes. The latter study is an example of how a school climate factor, under a different theoretical framework other than school effectiveness research, can address a non-scholastic outcome that is still of great relevance for educational aims.

[16] Thailand and Dominican Republic displayed inconsistent directions of effects for two outcomes (support for ethnic rights and for freedom of speech). We inspected the matrix of correlations, and these two countries presented the lowest correlations between authoritarianism with all outcomes, and also had the highest country means on authoritarianism. Taken altogether, these may suggest ceiling effects of the mediator.

Implications of the work

The contributions of the present work may be divided into general contributions and methodological contributions. The first set concerns the broad theoretical implications of the empirical papers for our understanding of school climate, as well as their practical relevance, whereas the second set concerns the methodological insights and innovations that have been generated by the analytic approaches undertaken in the thesis.

General contributions

Overall **theoretical contributions to school climate research.** There is an advantage in taking a non-exhaustivity approach to school climate factors. We feel that all-encompassing models of school climate share a central limitation, namely restricting the attention paid to school level constructs by focusing on school outcomes pertaining to school adjustment in general, mainly academic achievement. This categorization has little to offer with respect to illuminating non-scholastic outcomes, such as students' democratic values, and teachers' organizational behaviour. There is a need to open up the broad array of school climate domains, and specify to which outcomes the various elements of school climate are connected.

This distinction is not new. This approach was actually adopted earlier, within the organizational literature during the 1970s (Schneider, Ehrhart, & Macey, 2011). The dilemma was to have either: a) a single exhaustive list of domains on organizational climate, i.e. the molar climate approach to in general describe the work experience; or b) a distilled list of organizational climate factors for certain outcomes, i.e. the "climates for" approach. The latter led to concepts of a climate for safety, climate for quality service, and climate for innovation, to name a few. This kind of approach, however, needs considerable work before it can be seen as compelling and convincing. The present work is an example of how different school climate factors are specified as relevant "climates for" different outcomes, based on theory driven models. The details of our contribution through the four papers are specified below.

Broadening the scope of school climate research. Wang and Degol (2015) asserted that one of the main gaps in school climate research, was its principal focus on students' outcomes, and less attention to teachers' outcome. The present work was designed to fill this gap, illustrating applications of school climate research to teachers' turnover (actual behaviours and intentions), and to students' non-scholastic characteristics, such as democratic attitudes.

The study of teacher turnover is relevant for educational equity and public policy. Uneven allocation of teachers between schools partially explains unequal results between schools (Akiba et al., 2007). Moreover, in the case of Chile, teacher career trajectories deepen these effects (Rivero, 2015). In paper 1, we addressed this problem and provided population estimates of teacher turnover rates for first contract teachers. Our results complement previous research from Valenzuela and Sevilla (2013) which estimated overall turnover rates between 2000-2008. Taking a different approach, we provided survival estimates which give more information for identifying the riskiest profiles, in terms of teachers' characteristics, working conditions, and school differences, for the cohort of teachers followed from 2007 to 2013. Unlike R. Allen et. al (2012), who suggest age is one of the factors that explains higher attrition in disadvantaged schools in England, our results shows independent effects of school intake, school type, age, sex, and teachers' role. Within the same paper we evaluated the teachers' experience of school climate factors such as leadership and management priorities (staff relations and academic monitoring), and showed these factors are relevant protective factors for teacher turnover. Moreover, in paper 2 we used a representative sample of teachers from TALIS 2013 to thoroughly estimate the relative contribution of school climate factors, such as teacher-student relations and classroom discipline, and showed how these serve as protective factors for teacher intentions to leave the profession and intentions to turnover from schools. Taken together these two papers shed light on retention factors, including structural working conditions set by law (salaries, system of contracts), as well as less tangible factors relative to school management practices.

These results are relevant for public policy purposes, particular in connection with efforts to reduce teacher turnover rates, and indirectly to change the uneven distribution of teachers between schools. Chile presents high rates of attrition of novice teachers in comparison to other countries (Valenzuela & Sevilla, 2013), and a high proportion of senior teachers are also soon to retire in comparison to other countries (OECD, 2005). These two features will focus the educational system in future years on recruiting more teachers. Currently, a new law for the teaching career is in discussion in the Chilean parliament (MINEDUC, 2015). This would set new conditions of teachers' recruitment, preparation, evaluation of teachers, salaries, professional development, and working conditions. Preparing highly qualified teachers and ensuring the attraction of more qualified teachers aims to boost the stock of the system capacities. Yet, Chile has been characterized for a lack of policies to strengthen induction of new teachers in order to buffer against teacher turnover (Avalos & Aylwin, 2007; Avalos & Sevilla, 2013; Avalos, 2009). The new policy discussion is an opportunity to address the problem of teacher turnover.

It should be noted that the study of teacher turnover in particular is not so different to the study of job turnover in general. Thus, the application of work and organizational theories can be straightforward, while considering the particularities of the school system. This consideration demands more organizational approaches that shed light on organizational process and system level properties other than teacher characteristics and school socioeconomic differences (Ingersoll & Smith, 2003; Ingersoll, 2001). If such approaches are less prominent in relative terms within the literature, they are even more scarce for the study of the teachers in Chile in particular. The present work, alongside that of others (e.g., Avalos & Sevilla, 2013; Rivero, 2013, 2015; Valenzuela, 2013), has the potential to change this situation by drawing attention to the school climate factors that play a role in teacher turnover behaviours and intentions.

The centrality of students' achievement within school climate research often overshadows other, non-scholastic covariates and outcomes. This is a research gap less noted in school climate research reviews. In paper 3, we modeled the indirect effects of bullying on achievement, because understanding the pathways by which these factors are related may shed light on how to intervene in such a process. Our results are consistent with the idea of 'attunement' from Norwalk and colleagues (2015). The authors suggested that when teachers' perceptions of bullying events (who bullies who) are in alignment with those of the students, students' sense of belonging is higher and students' peers are less tolerant of bullying and behave in a protective manner towards victims. While paper 3 did not focus on concordance in perceptions of bullying, it does have implications for how the experience of the broader school climate (in this case, feelings of school belonging) may moderate the impact of bullying that has already occurred. If the process by which bullying affects students' engagement and indirectly achievement is interrupted, then interventions towards bullying could potentially work not only by targeting the antecedents of bullying (preventive actions to reduce bullying) but also by an ongoing corrective intervention at the level of school belonging in order to reduce negative effects of bullying. Highlighting the contextual effects of bullying rates and schools' levels of belonging may also shed light on how to identify school difficulties in this important area of children's adjustment at school.

Bullying is a pervasive phenomenon in Chilean schools, as we argued earlier. In 2011, Chile a law on school violence was promulgated (MINEDUC, 2011a), accompanied by the "Plan Escuela Segura" (Safety School plan) (MINEDUC, 2012). This law introduced a process of monitoring violent events within schools, demanded that schools introduce administrative procedures to handle violent events, and enforced accountability measures to promote adherence to school rules. The high emphasis on accountability measures has been criticised by Magendzo, Toledo, and Gutiérrez (2012), rooted in a paradigm of control and sanction according to the authors. This programme contrasts with earlier programmes from the late 1990s such as "Habilidades para la vida" (Skills for Life) from "Junta de Auxilio Escolar y Becas" (JUNAEB) (Madriaza, 2008), which were designed to foster communal values and a safety environment for all students. The results from paper 3 provide evidence more consistent with these latter approaches, by unravelling the role of belonging and engagement as mediators of the effects of bullying for achievement. However, it should be acknowledged that bullying presents more worrisome consequences than academic achievement, such as overall wellbeing, health, and risk of suicides (UNESCO, 2013). The present work is therefore a small contribution towards understanding and intervening in a much bigger problem.

In paper 4, we addressed the role of open classroom discussion in students' democratic attitudes endorsement. These are perhaps the most unconventional outcomes included in the present piece of work in comparison to previous studies. Nevertheless, as argued throughout this thesis, any outcome that is subject to clustering may be worthy of school climate research inquiry, if clustered level effects do exist and if the analysis is rooted in a relevant conceptual model. This also has the advantage of addressing the highly achievement-centric focus on school climate research, although this requires innovative analyses of the connection between democratic attitudes and the school context. During a presentation of this paper at the International Research Conference (Carrasco & Banerjee, 2015b), Bruno Losito – one of the principal investigators of the International Civic and Citizenship Education Study (ICCS 2009) – indicated that one of the main problems to resolve is what we can recommend to educators with this study. This particular paper offers a distinctive contribution by unravelling the role that teachers may play in creating classrooms open for discussion.

Teachers can foster an open classroom for discussion by: encouraging students to make up their minds and express their opinions, encouraging students to discuss different points of view, and explaining different sides of contentious issues. With this practice, teachers foster the endorsement of democratic attitudes, such as gender equality and the condemnation of corruption. Eleven years have passed since the "Comision Formación Ciudadana" (Commission for Citizenship Formation) (MINEDUC, 2004b) and the 2004 curricular reform for citizenship in Chile (MINEDUC, 2004a). In this reform, civic education, conceived as knowledge of the state and political system, was considered unsatisfactory for the full development of citizens. Thus, the concept of "formación ciudadana" (citizenry development) was introduced, designed to guarantee the scope of knowledge, abilities and attitudes for the civic life. The inclusion of the dimension of attitudes as an educational objective is a new development for this context. Current curricular analysis shows that Chilean curricula still present a stronger emphasis on the formal side of citizenship: the knowledge of institutions, the rule of law and patriotism (Bascopé et al., 2013), casting doubt on how successfully the aims of fostering democratic dispositions are fulfilled, and what actual practices within school serve that purpose. Our work shows that open classroom discussion operates as a moderator that appears to boost the relation between civic knowledge and general democratic values – support for gender equality, ethnic rights, and freedom of speech, and the condemnation of corruption – all of which are important dispositions for citizenry in democratic countries (Torney-Purta, Amadeo, & Pilotti, 2004).

Methodological contributions

Our conceptualisation of school climate as a multifaceted concept linked to a variety of school outcomes, beyond academic achievement, for both teachers and students, raised significant methodological and analytic challenges. The nested nature of the observations of teachers' and students' outcomes needed to be addressed by appropriate partition of the within and between parts of covariates under study. In the current work, we used both multilevel models and clustered error models to do this. The inferential level was held at the school for all covariates and outcomes when possible.

The only study in which this was not possible for school climate factors was in paper 1, due to the sparseness of the data, in which schools were represented by too few teachers to create valid clustered level aggregations. However, in all other papers, the referent nature of the indicators of each construct was considered, in order to specify appropriate models and to produce valid estimates.

Model specifications are ways to formalize our conceptual models, i.e. our expectations of relations and effects between variables. The model we fitted in each paper can therefore be viewed as a formalization of what we think is relevant for the problem at hand. Given our consideration of mediated and moderated associations among variables, a number of different types of complex models were specified, enabling the estimation of indirect effects and conditional indirect effects. These effects are not straightforward to estimate with traditional approaches, such as regression and multilevel models, but they do raise challenges in terms of both carrying out and communicating appropriate analyses.

Appropriate data analysis for complex sample data. Within large scale assessments, one of the difficulties for new researchers is to use appropriate methods of estimation, given the special features these data have, such as sampling weights and plausible values (L. Rutkowski et al., 2010). There is a temptation to assume that model based methods (multilevel models) are sufficient to account for the complexities of these data sources. However, these approaches may not adequately account for information regarding the stratification of the sample and the sampling weights, distorting the inferences to the population. Sample design is indeed informative (Asparouhov, 2006), and only under exceptional conditions of correct model specification will estimates not be biased (Snijders & Bosker, 2012). Thanks to the contemporary literature on sampling weights, model specification for complex sample data and extensive documentation (e.g., Asparouhov & Muthén, 2004, 2006; Asparouhov, 2008; L. Rutkowski, Davier, & Rutkowski, 2013; Snijders & Bosker, 2012; Stapleton, 2002, 2006, 2013), the current work was able to make use of sampling weights, stratification information and plausible values methods. Thorough descriptions of the method of estimations are given in each of the papers, especially in papers 2 (population estimates with Balanced Repeated Replication and multilevel SEM with sampling weights), 3 (population estimates with jackknife variance estimation, multilevel with full sampling design, and multilevel SEM with sampling design), and 4 (SEM with Taylor Linearization), in order to aid further research with these data sources. Most of the time, the limitations are due to the use of software that is unable to provide appropriate methods for estimation that account for the full sampling design. For example, STATA cannot accommodate the stratification information of the sampling design, in multilevel models (West, Welch, & Galecki, 2014). Moreover, there is also often a lack of documented examples to help researchers to set up these methods of estimation. Kim and colleagues (2013) chapter, is good exception. However, full examples with data and code, are scarce. One of the aims of the current work was to bridge that gap, by thorough descriptions of the methods used in each paper. Further work needs to be done to make the current results public and replicable, by making code and data available for open inquiry.

Limitations and further directions

The current work has various limitations common to other studies on school climate research that have relied on survey methods in cross sectional designs. Results from this work are bounded by issues of causality, temporality, and generalizability. These limitations were briefly acknowledged and described in chapter 2. Below, these limitations are addressed in relation to the current work, and suggestions for how these can be addressed in future research are made.

Causality. Most of the indirect effects models (mediations) assume a theoretical casual direction of effects. However, as long as mediators and outcomes are contaminated by endogeneity, that is correlated disturbances due to unobserved common factors, there is no warrant for interpreting a causal effect (Antonakis et al., 2010). This is a pervasive problem with partial mediation effects, and most worrisome for models with fewer variables (Luchman, 2014). This problem is general to the current work, particularly for papers 2, 3 and 4, which present indirect effects models.

A further line of inquiry for the present work is to implement sensitivity analysis under the causal framework for mediation (Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010). This framework helps to assess the limits of model estimates, under violations of the sequential ignorability assumption. This assumption, in plain terms, states that the predictor x is orthogonal to the outcome and mediator, and the mediator is independent of the outcome. By evaluating how the model estimates hold under different scenarios of violations of the sequential ignorability assumption, the approach permits to illustrate if the indirect effects are tenable and robust for causal interpretation. There are current examples of this approach with large scale assessment (e.g., Caro, 2014; Ponzo, 2013). However, I have not found any applications of this framework to multilevel mediation models. Thus, further research on this line of inquiry would be worthwhile for school climate research development.

Ponzo (2013) used the propensity score matching technique to isolate students with similar characteristics, and then to estimate the effect of bullying on achievement within a potential outcome framework. As argued earlier, this should not be different to asking what would be the expected outcome for student *i*, if s/he attended school *j*, given that this school has a given set of clustered characteristics. This is applicable for all presented studies in the current work: what is the potential outcome for teacher turnover as a function of varying levels of school supportive leadership? what is the potential outcome for students' achievement as a function of varying levels of school sense of belonging? and what is the potential outcome for students' support for gender equality as a function of varying levels of school means of authoritarianism?

There is no reason to fall into the trap of "correlation is not causation" without careful reflection (Antonakis & Lalive, 2011). One of the main aims of research is to identify causes and effects, and provide explanations. While non-experimental data are frequently considered to only *suggest* possible causal relations, it is also the case that, under certain conditions (e.g., where propensity matching is used to compare outcomes for individuals who differ in a key predictor variable but who are matched on an array of relevant covariates), causal interpretation of effects can be undertaken even within a correlational design. Indeed, in some cases where randomisation of individuals to certain treatment experiences is not feasible or ethical (e.g., being bullied), this kind of work is essential for evaluating causal effects. The potential outcome framework of causal inference dwells on this challenge (Robinson, 2014). The present work made some strides forward in this regard, by including a range of individual and school covariates plausibly associated with the key school climate predictors, but future work could use more extensive propensity matching techniques to strengthen the analysis. Large scale assessment studies present advantages to this task, thanks to their sample size, multi country samples, random selection of observations, and public access to data.

Temporality. The issue of causality is linked to important questions about temporality. Especially in the absence of adequate matching of individuals on relevant covariates, cross-sectional datasets raise the challenge of whether effects might be occurring in an opposite direction, or even operating in a bidirectional manner. The use of longitudinal research designs would serve to assess the directional effects assumptions made in most of the mediational models outlined in the present manuscript. Cross-lagged designs can suit this purpose, but, given the cross sectional nature of the large-scale assessments used in papers 2, 3 and 4, the current work has little to offer for shedding light on the role of school climate within longitudinal trajectories. Indeed, work on longitudinal effects of school climate is still in its infancy. This is a research gap noted in Anderson (1982) in 1980's, which is still valid (Wang & Degol, 2015). This is therefore an important area for further research.

In fact, the study of school climate factors and clustered level effects within longitudinal designs is a general research gap within this area of research. This is of special importance for the study of outcomes that by their nature manifest themselves over time. Within the current work, this is especially the case for the study of teacher turnover. Only study 1 from paper 1 had a longitudinal design. Yet we did not explore fully the potential effects of variations in factors across years. To fully study the effects of school climate factors on censored outcomes such as teacher turnover and students' school dropout, among other events, an additional issue needs to be considered. This is the problem of non-ignorable missing data patterns, where the covariates are related to the attrition process (Enders, 2011). Let us say that we follow novice teachers for four years in order to evaluate an induction program. School climate factors such as teacher-student relations and head teacher support are measured over time. The aim is to capture variations of school climate over time, jointly with changes in the turnover rates. Because the outcome entails teachers leaving the cluster of interest, how would the teachers in the 'non-stayer' condition (the teachers who left the school) provide information about the school environment that they left? Moreover, the missing at random mechanism does not hold, because low levels of the school climate factors are expected to predict turnover. In such scenarios, researchers need to deal with this special case of missing data, namely monotonic missing patterns (Schafer & Graham, 2002), with missing not at random (Enders, 2011). This issue is not addressed in the school climate literature reviewed in this manuscript. Examples of this line of inquiry, in which covariates vary over time, do however exist. Muthén and Masyn (2005) used such an approach to estimate the relations between aggressive trajectories and school removal. Similarly, Lamote and Damme (2013) conducted research on school dropout using a multilevel discrete model, accounting for school changes, and accounting for school level effects. These two studies address the issue not only that the outcome changes over time, but also that time-variant processes operate for the other covariates. Similar models could be applied for the study or teacher turnover, while addressing the missing data problem, with selection models or pattern mixture (Enders, 2011). However, examples of this kind of approach with clustered level variations of school climate factors have yet to be found.

Generalizability. Results in papers 1, 2 and 3 were restricted to Chilean samples. Although the sample in study 1 of paper 1 was at the census level, these results may not generalize to other contexts. Study 2 in paper 1 was even more limited, because the National Teacher Evaluation System recruiting process does not follow a random sampling strategy. In spite of the large samples it recruits at each cohort, they have diverse characteristics from one year to the next. One way to strength the results of this study in particular, is to replicate the model-testing with other cohorts that share the same measures. Even if this is attempted, results will need to be interpreted with caution, because the selection process of the sample is not random, and strict covariate controls should be used.

Results in paper 2 came from a representative sample of lower secondary teachers from Chile who participated in TALIS 2013. The results presented in this papers are generalizable only to this sampling frame. Therefore, these results cannot be simply assumed to generalize to primary school teachers or to teachers that teach other age groups. However, the generalizability of the results can be fostered by replicating the present model with data from other countries that participated in TALIS 2013. This endeavour can clarify how sensitive the model is to cultural variations, economic differences between countries, and educational system level properties.

Paper 3 also used a random sample from Chile. These were students from secondary schools with a mean age of 14 years. Thus, the results are generalizable to this population, but not to other age cohorts and other contexts. A straightforward way to gain more generalizability for these model would be to fit the same model to data from the younger sample of students that participated in TIMSS 2011 and to data from more countries.

Paper 4 is the only paper which used a multi-country sample to assess the effects of interest. In general, this approach gives more evidence in favour of the presented model. However, because the variables were not the same across all countries for authoritarianism and for one of the outcomes, the generalizability of results is limited to the conceptual analysis of effects observed when comparing results from Latin American countries and Asian countries.

A further line of development consists of evaluating the measurement models of the current measures to provide assurance that these effects are comparable beyond the mere direction of effects, thus including also the valid interpretation of the mean levels between countries. Even if the comparison between Asia and Latin America will be limited, this kind of endeavour will strength the generalizability of results by providing comparable estimates for the countries of each region. This kind of evidence of measurement invariance between measures across countries would be helpful for papers 2 and 3 too. This can include test of invariance for countries, as well for the inferential levels of interest in the study (in this case, the school level). This kind of work would confirm that the multiple indicator scales present the same measurement properties for individuals within clusters, between schools, and across countries. Doubly latent models achieve such a design (Lüdtke et al., 2011; Marsh et al., 2009; Morin et al., 2014).

Additional school outcomes. The present work has gone beyond students' achievement to demonstrate the broad relevance of school climate to a variety of school outcomes. Yet there are, of course, a wide variety of other outcomes that deserve similar attention. These include health-related behaviours of students (e.g., substance use, aggression, sexual activity behaviours), political socialization (e.g., prejudice, democratic values, political participation, political sophistication), and non-traditional skills (e.g., digital literacy, financial literacy), among others. The same can be said for teachers, on health-related outcomes (e.g., stress, burnout, depression), work and organizational behaviours (e.g., performance, commitment, organizational citizenship behaviours, withdrawal behaviours, turnover, attrition), and other non-traditional skills. The research reported here thus sets an agenda for a more expansive consideration of the complex array of outcomes known to be linked to teachers' and students' experience at school (Wang & Degol, 2015).

Final Conclusion

The overall aim of the current research was to address the relations between school climate and school outcomes. Given the abstract nature of this question, I firstly reviewed the evidence to arrive at a stance on how best to conceptualize school climate. School climate has been portrayed in the present thesis as an umbrella concept for various factors that describe cluster-level differences in the school environment, either at the classroom level or at the school level as a whole. Extending far beyond the achievement-centric focus of much of the existing literature, we investigated a wide range of relevant educational outcomes besides student achievement, such as teacher turnover (actual behaviour and intentions) and students' democratic attitudes.

Throughout this work, a multi-faceted school climate factors approach is proposed, rejecting a uni-dimensional school climate construct, and instead adopting a "climate for" approach in relation to different outcomes. This extends school effects methods through use of multivariate approaches for different scenarios (e.g., survival outcomes, binary outcomes, multilevel structures, clustered approaches to clustered effects). A great emphasis of the current works dwells on "how to answer" school climate research questions, more than to exhaustively pursue a question about a single construct within one theoretical framework. This latter feature of the thesis explains the diverse outcomes and theories presented in the current work. I hope that the empirical work fosters further development in school climate research and provides a theoretical and methodological foundation for addressing further relevant complex questions.

References

Adams, G. J. (1996). Using a Cox Regression Model Examine Teacher Turnover. *The Journal of Experimental Education*, *64*(3), 267–285.

Adorno, B. T. W., Frenkel-brunsivik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper & Row.

Agencia de Calidad de la Educación. (2011). *Resultados TIMSS 2011 Chile: Estudio Internacional de Tendencias en Matemática y Ciencias*. Santiago de Chile: Ministerio de Educación de Chile & International Association for the Evaluation of Educational Achievement. Retrieved from https://s3.amazonaws.com/archivos.agenciaeducacion.cl/documentos-web/Estudios+Internacionales/TIMSS/Informe+Nacional+de+Resultados+TIMSS+2011.pdf

Agencia de Calidad de la Educación. (2012a). *Informe Técnico Simce 2012*. Santiago, Chile: Agencia de Calidad de la Educación. Retrieved from https://s3-us-west-2.amazonaws.com/documentos-web/Informe_Tecnico_Simce_2012.pdf

Agencia de Calidad de la Educación. (2012b). Metodología de Construcción de grupos Socioeconómicos Pruebas SIMCE 2012.

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher Quality, Opportunity Gap, and National Achievement in 46 Countries. *Educational Researcher*, *36*(7), 369–387. http://doi.org/10.3102/0013189X07308739

Alfonso, M., Santiago, A., & Bassi, M. (2010). An Alternative Pathway Into Teaching: Placing Top University Graduates in Vulnerable Schools in Chile, (February). Retrieved from http://services.iadb.org/wmsfiles/products/Publications/35130716.pdf

Algina, J., & Swaminathan, H. (2011). Centering in two-level nested designs. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 285–312). Taylor & Francis Group.

Allen, M. B. (2005). *Eight Questions on Teacher Recruitment and Retention: What Does the Research Say? Education commission of the States*. Denver, CO, US. Retrieved from http://files.eric.ed.gov/fulltext/ED489332.pdf

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, *63*(1), 1–18. http://doi.org/10.1111/j.2044-8325.1990.tb00506.x

Allen, R. (2013). The recruitment and retention of teachers in deprived schools [Slides]. Retrieved from http://sticerd.lse.ac.uk/seminarpapers/ses20022013.pdf

Allen, R., & Burgess, S. (2010). How long are teachers staying? Bristol, UK: Centre for Market and Public Orgazanization. Retrieved from http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/howlongteachersstaying.pdf

Allen, R., Burgess, S., & Mayo, J. (2012). The teacher labour market , teacher turnover and disadvantaged schools : new evidence for England. Bristol, UK: Centre for Market and Public Organisation. Retrieved from http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp294.pdf

Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009). *The Schools Teachers Leave. Teacher Mobility in Chicago Public Schools*. Consortium on Chicago School Research, at the University of Chicago Urban Education Institute. Retrieved from http://ccsr.uchicago.edu/sites/default/files/publications/CCSR_Teacher_Mobility.pdf

Altemeyer, B. (1981). *Right-wing authoritarianism*. University of Manitoba Press.

Anderson, C. S. (1982). The Search for School Climate: A Review of the Research. *Review of Educational Research*, *52*(3), 368–420. http://doi.org/10.3102/00346543052003368

Anderson, H. H., & Brewer, H. M. (1945). *Studies of Teachers' Classroom Personalities, I: Dominative and Socially Integrative Behavior of Kindergarten Teachers*. London, Humphrey Milford: Stanford University Press.

Anderson, H. H., & Brewer, J. E. (1946). *Studies of Teachers' Classroom Personalities, II. Effects of teachers' dominative and integrative contacts on children's classroom behavior.* Stanford University Press.

Anderson, H. H., Brewer, J. E., & Reed, M. F. (1946). *Studies of Teachers' Classroom Personalities, III. Follow-up studies of the effects of dominative and integrative contacts on children's behavior.* Stanford University Press.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. http://doi.org/10.1016/j.leaqua.2010.10.010

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2014). Causality and endogeneity: Problems and solutions. In D. V. Day (Ed.), *The Oxford Handbook of Leadership and Organizations.* (pp. 93–117). New York: New York: Oxford University Press.

Antonakis, J., & Lalive, R. (2011). Counterfactuals and Causal Inference: Methods and Principles for Social Research By Stephen L. Morgan & Christopher Winship. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(1), 152–159. http://doi.org/10.1080/10705511.2011.532697

Argyris, C. (1958). Some problems in conceptualizing organizational climate: A case study of a bank. *Administrative Science Quarterly*, 501–520.

Asbrock, F., Christ, O., Duckitt, J., & Sibley, C. G. (2012). Differential effects of intergroup contact for authoritarians and social dominators: a dual process model perspective. *Personality & Social Psychology Bulletin*, *38*(4), 477–90. http://doi.org/10.1177/0146167211429747

Asbrock, F., Sibley, C. G., & Duckitt, J. (2010). Right-wing authoritarianism and social dominance orientation and the dimensions of generalized prejudice: A longitudinal test. *European Journal of Personality*, *24*(4), 324–340. http://doi.org/10.1002/per.746

Ashkanasy, N. M., Wilderom, C. P. M., & Peterson, M. F. (2001). Handbook of Organizational Culture & Climate. *Human Resource Management*, *40*(2), 189–191.

Asparouhov, T. (2006). General Multi-Level Modeling with Sampling Weights. *Communications in Statistics - Theory and Methods*, *35*(3), 439–460.

http://doi.org/10.1080/03610920500476598

Asparouhov, T. (2008). Scaling of sampling weights for two level models in Mplus 4.2. Los Angeles, CA: Muthén & Muthén. Retrieved from https://www.statmodel.com/download/Scaling3.pdf

Asparouhov, T., & Muthén, B. (2004). Weighting for unequal probability of selection in latent variable modeling. Los Angeles, CA: Muthén & Muthén. Retrieved from https://www.statmodel.com/download/webnotes/MplusNote81.pdf

Asparouhov, T., & Muthén, B. (2005). Multivariate Statistical Modeling with Survey Data. Los Angeles, CA: Muthén & Muthén. Retrieved from http://statmodel2.com/download/AsparouhovMuthen_MultivariateModeling3.pdf

Asparouhov, T., & Muthén, B. (2006). Multilevel modelling of complex survey data. In *Joint Statistical Meeting, ASA section on Survey Research Methods* (pp. 2718–2726). Seattle, WA. Retrieved from http://www.statmodel.com/download/SurveyJSM1.pdf

Asparouhov, T., & Muthén, B. (2010). Resampling Methods in Mplus for Complex Survey Data. *Mplus Technical Report*. Los Angeles, CA. Retrieved from https://www.statmodel.com/download/Resampling_Methods5.pdf

Avalos, B. (2009). La inserción profesional de los docentes. *Profesorado*, *13*(1), 43–59. Retrieved from https://www.ugr.es/~recfpro/rev131ART3.pdf

Avalos, B., & Aylwin, P. (2007). How young teachers experience their professional work in Chile. *Teaching and Teacher Education*, *23*(4), 515–528. http://doi.org/10.1016/j.tate.2006.11.003

Avalos, B., & Sevilla, A. (2013). Teaching and Learning "For All": Teachers and Future Teachers in Chile. In *UKFIET International Conference on Education and Development – Education & Development Post 2015: Reflecting, Reviewing and Re-visioning*. Oxford, UK.

Azeredo, C. M., Rinaldi, A. E. M., de Moraes, C. L., Levy, R. B., & Menezes, P. R. (2015). School bullying: A systematic review of contextual-level risk factors in observational studies. *Aggression and Violent Behavior*, *22*, 65–76. http://doi.org/10.1016/j.avb.2015.04.006

Bakker, A. B., & Bal, M. P. (2010). Weekly work engagement and performance: A study among starting teachers. *Journal of Occupational and Organizational Psychology*, *83*(1), 189–206. http://doi.org/10.1348/096317909X402596

Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: state of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. http://doi.org/10.1108/02683940710733115

Bakker, A. B., Hakanen, J. J., Demerouti, E., & Xanthopoulou, D. (2007). Job resources boost work engagement, particularly when job demands are high. *Journal of Educational Psychology*, *99*(2), 274–284. http://doi.org/10.1037/0022-0663.99.2.274

Banerjee, R., Weare, K., & Farr, W. (2014). Working with "Social and Emotional Aspects of Learning" (SEAL): Associations with school ethos, pupil social experiences, attendance, and attainment. *British Educational Research Journal*, *40*(4), 718–742. http://doi.org/10.1002/berj.3114

Barker, R. G. (1963). On the nature of the environment. *Journal of Social Issues*, *19*(4), 17–38.

Barker, R. G., & Gump, P. V. (1964). *Big School, Small School: High School Size & Student Behavior*. Stanford University Press.

Bascopé, M., Bonhomme, M., Cox, C., Castillo, J. C., & Miranda, D. (2013). National Curricular Guidelines and Citizenship Education in Schools in Latin American Countries. In *5th IEA International Research Conference*. Singapore. Retrieved from http://www.iea.nl/fileadmin/user_upload/IRC/IRC_2013/Papers/IRC-2013_Bascope_etal.pdf

Bauman, S., Toomey, R. B., & Walker, J. L. (2013). Associations among bullying, cyberbullying, and suicide in high school students. *Journal of Adolescence*, *36*(2), 341–350. http://doi.org/10.1016/j.adolescence.2012.12.001

Begg, M. D., & Parides, M. K. (2003). Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Statistics in Medicine*, *22*(16), 2591–602. http://doi.org/10.1002/sim.1524

Bell, B. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010). The Impact of Small Cluster Size on Multilevel Models : A Monte Carlo Examination of Two-Level Models with Binary and Continuous Predictors. In *JSM Proceedings, Survey Research Methods Section* (pp. 4057–4067). Retrieved from http://www.amstat.org/sections/srms/Proceedings/y2010/Files/308112_60089.pdf

Bellei, C., & Valenzuela, J. P. (2010). ¿Están las condiciones para que la docencia sea una profesión de alto estatus en Chile? In S. Martinic & G. Elacqua (Eds.), *¿Fin de Ciclo? Cambios en la Gobernanza del Sistema Educativo*. Santiago, Chile: Facultad de Educación, Pontificia Universidad Católica de Chile y Oficina Regional para América Latina y el Caribe UNESCO. Retrieved from http://www.researchgate.net/publication/273143530

Benner, A. D. (2000). The cost of teacher turnover. Austin, Texas: Texas Center for Educational Research. Retrieved from http://gse.berkeley.edu/sites/default/files/users/bruce-fuller/LAUSD-Berkeley ReedSchoolClimate_TeacherEngagementAssessment_ TechnicalReport_Dec2013.pdf

Berger, C., Rodkin, P. C., & Karimpou, R. (2008). Bullies and victims at school: perspectives and strategies for primary prevention. In T. Miller (Ed.), *School Violence and Primary Prevention* (pp. 287–314). New York: Springer-Verlag.

Berry, C. M., Lelchook, A. M., & Clark, M. A. (2012). A meta-analysis of the interrelationships between employee lateness, absenteeism, and turnover: Implications for models of withdrawal behavior. *Journal of Organizational Behavior*, *33*(5), 678–699. http://doi.org/10.1002/job.778

Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The Effects of Educational Accountability on Teachers: Are Policies Too Stress Provoking for Their Own Good? *International Journal of Education Policy and Leadership*, *4*(5). Retrieved from http://journals.sfu.ca/ijepl/index.php/ijepl/article/view/99

Billig, M. (2013). *Learn to Write Badly: How to Succeed in the Social Sciences*. Cambridge University Press.

Billingsley, B. S. (2004). Special Education Teacher Retention and Attrition: A Critical Analysis of the Research Literature. *The Journal of Special Education*, *38*(1), 39–55. http://doi.org/10.1177/00224669040380010401

Blalock, H. M. (1984). Contextual-Effects Models: Theoretical and Methodological Issues. *Annual Review of Sociology*, *10*(1), 353–372. http://doi.org/10.1146/annurev.so.10.080184.002033

Blau, G., & Lunz, M. (1998). Testing the Incremental Effect of Professional Commitment on Intent to Leave One's Profession beyond the Effects of External, Personal, and Work-Related Variables. *Journal of Vocational Behavior*, *52*(2), 260–269. http://doi.org/10.1006/jvbe.1997.1601

Boe, E. E., Bobbitt, S. a., & Cook, L. H. (1997). Whither Didst Thou Go? Retention, Reassignment, Migration, and Attrition of Special and General Education Teachers from a National Perspective. *The Journal of Special Education*, *30*(4), 371–389. http://doi.org/10.1177/002246699703000402

Bonett, D. G. (2007). Transforming odds ratios into correlations for meta-analytic research. *American Psychologist*, *62*(3), 254–255. http://doi.org/10.1037/0003-066X.62.3.254

Borman, G. D., & Dowling, N. M. (2008). Teacher Attrition and Retention: A Meta-Analytic and Narrative Review of the Research. *Review of Educational Research*, *78*(3), 367–409. http://doi.org/10.3102/0034654308321455

Bothma, C. F. C., & Roodt, G. (2013). The validation of the turnover intention scale. *SA Journal of Human Resource Management*, *11*(1), 1–13. http://doi.org/10.4102/sajhrm.v11i1.507

Bovard, E. W. (1951). The psychology of classroom interaction. *The Journal of Educational Research*, *45*(3), 215–224. Retrieved from http://www.jstor.org/stable/27529355

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2010). The Influence of School Administrators on Teacher Retention Decisions. *American Educational Research Journal*, *48*(2), 303–333. http://doi.org/10.3102/0002831210380788

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *American Economic Review*, *95*(2), 166–171. http://doi.org/10.1257/000282805774669628

Bravo, D., Urrutia, D., & Peirano, C. (2006). *Encuesta Longitudinal de Docentes 2005 : Análisis y Principales Resultados*. Retrieved from http://ww2.educarchile.cl/UserFiles/P0001/Image/Documento ELD Final.pdf

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2011a). *ICCS 2009 User Guide for the International Database*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2011b). *ICCS 2009 User Guide for the International Database Supplement 1*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2011c). *ICCS 2009*

User Guide for the International Database, Supplement 3. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Brookover, W. B. (1979). *School social systems and student achievement: schools can make a difference*. New York: Praeger.

Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. (1978). Elementary School Social Climate and School Achievement. *American Educational Research Journal*, *15*(2), 301 –318. http://doi.org/10.3102/00028312015002301

Brown, R. (2010). *Prejudice: Its Social Psychology* (2nd ed.). Chichester, UK: Wiley-Blackwell.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York and London: Guilford Press.

Bryk, A. S., & Schneider, B. L. (2004). *Trust In Schools: A Core Resource For Improvement*. New York: Russell Sage Foundation.

Bryk, A. S., Sebring, P. B., & Allensworth, E. (2010). *Organizing schools for improvement: lessons from Chicago*. Chicago and London: University of Chicago Press.

Buckley, J., Schneider, M., & Shang, Y. (2005). Fix It and They Might Stay: School Facility Quality and Teacher Retention in Washington, D.C. *Teachers College Record*, *107*(5), 1107–1123. http://doi.org/10.1111/j.1467-9620.2005.00506.x

Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology*, *98*(1), 1–13. http://doi.org/10.1037/0022-0663.98.1.1

Buhs, E. S., Ladd, G. W., & Herald-Brown, S. L. (2009). Victimization and Exclusion: Links to Peer Rejection, Classroom Engagement, and Achievement. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of Bullying in Schools: An International Perspective* (pp. 163–172). Routledge.

Burkhouse, K. L. S. (2009). *The Measurement of School Climate Using Surveys: Exploring Unit of Analysis*. University of Maryland. Retrieved from http://drum.lib.umd.edu/handle/1903/9987

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education*, *8*(1), 158–233. http://doi.org/10.3102/0091732X008001158

Buswell, M. M. (1953). The Relationship between the Social Structure of the Classroom and the Academic Success of the Pupils. *The Journal of Experimental Education*, *22*(1), 37–52. http://doi.org/10.2307/20153914

Cabezas, V., Gallego, F., Santelices, V., & Zarhi, M. (2011). *Factores Correlacionados con las Trayectorias Laborales de Docentes en Chile, con Especial Enfasis en sus Atributos Académicos*. *Proyecto FONIDE N°: FS511082- 2010*. Santiago, Chile. Retrieved from http://sgdce.mineduc.cl/descargar.php?id_doc=201208211511240

Campbell, D. E. (2008). Voice in the Classroom: How an Open Classroom Climate

Fosters Political Engagement Among Adolescents. *Political Behavior*, *30*(4), 437–454. http://doi.org/10.1007/s11109-008-9063-z

Canetti-Nisim, D. (2004). The effect of religiosity on endorsement of democratic values: The mediating influence of authoritarianism. *Political Behavior*, *26*(4), 377–398. http://doi.org/10.1007/s11109-004-0901-3

Carmody, S. (2014). *ngramr: Retrieve and plot Google n-gram data*. R package version 1.4.5. Retrieved from https://cran.r-project.org/web/packages/ngramr/ngramr.pdf

Caro, D. (2014). Causal mediation in educational research : An illustration using international assessment data. In *EARLI SIG Educational Effectiveness Conference* (pp. 1–29). Southampton, UK. Retrieved from https://www.researchgate.net/profile/Daniel_Caro4/publication/275655845_Causal_mediation_in_educational_research/links/554380850cf24107d3962e8b.pdf

Caro, D., & Lenkeit, J. (2012). An analytical approach to study educational inequalities: 10 hypothesis tests in PIRLS 2006. *International Journal of Research & Method in Education*, *35*(1), 3–30. http://doi.org/10.1080/1743727X.2012.666718

Caro, D., & Mirazchiyski, P. (2012). Socioeconomic gradients in eastern european countries: Evidence from PIRLS 2006. *European Educational Research Journal*, *11*(1), 96–110. http://doi.org/10.2304/eerj.2012.11.1.96

Caro, D., & Schulz, W. (2012). Ten Hypotheses about Tolerance toward Minorities among Latin American Adolescents. *Citizenship, Social and Economics Education*, *11*(3), 213–234. http://doi.org/10.2304/csee.2012.11.3.213

Carr, J. Z., Schmidt, A. M., Ford, J. K., & DeShon, R. P. (2003). Climate perceptions matter: A meta-analytic path analysis relating molar climate, cognitive and affective states, and individual level work outcomes. *Journal of Applied Psychology*, *88*(4), 605–619. http://doi.org/10.1037/0021-9010.88.4.605

Carrasco, D., & Banerjee, R. (2015a). Teacher Turnover: an organizational account of public schools. In *3er Encuentro Anual de la Red de Investigadores Chilenos en Educación (RED ICE)*. Sheffield, UK.

Carrasco, D., & Banerjee, R. (2015b). Walk the talk! Civic knowledge, democratic values, and the role of school climate for open discussion: A moderated mediation approach. In *IEA International Research Conference 2015* (pp. 1–31). Cape Town, South Africa.

Carsten, J. M., & Spector, P. E. (1987). Unemployment, job satisfaction, and employee turnover: A meta-analytic test of the Muchinsky model. *Journal of Applied Psychology*, *72*(3), 374–381. http://doi.org/10.1037/0021-9010.72.3.374

Carvacho, H., Zick, A., Haye, A., González, R., Manzi, J., Kocik, C., & Bertl, M. (2013). On the relation between social class and prejudice: The roles of education, income, and ideological attitudes. *European Journal of Social Psychology*, *43*(4), 272–285. http://doi.org/10.1002/ejsp.1961

Cerit, Y. (2009). The Effects of Servant Leadership Behaviours of School Principals on Teachers’ Job Satisfaction. *Educational Management Administration & Leadership*, *37*(5), 600–623. http://doi.org/10.1177/1741143209339650

Cha, S.-H. (2008). *Explaining teachers’ job satisfaction, intent to leave, and actual turnover: A structural equation modeling approach.* Florida State University.

Retrieved from http://diginole.lib.fsu.edu/islandora/object/fsu:182180/datastream/PDF/view

Chávez, R. C. (1984). The Use of High-Inference Measures To Study Classroom Climates: A Review. *Review of Educational Research*, *54*(2), 237–261. http://doi.org/10.3102/00346543054002237

Christopher, A. N., & Wojda, M. R. (2008). Social Dominance Orientation, Right-Wing Authoritarianism, Sexism, and Prejudice Toward Women in the Workforce. *Psychology of Women Quarterly*, *32*(1), 65–73. http://doi.org/10.1111/j.1471-6402.2007.00407.x

Chudgar, A., & Luschei, T. F. (2009). National Income, Income Inequality, and the Importance of Schools: A Hierarchical Cross-National Comparison. *American Educational Research Journal*, *46*(3), 626–658. http://doi.org/10.2307/40284857

Ciani, K. D., Middleton, M. J., Summers, J. J., & Sheldon, K. M. (2010). Buffering against performance classroom goal structures: The importance of autonomy support and classroom community. *Contemporary Educational Psychology*, *35*(1), 88–99. http://doi.org/10.1016/j.cedpsych.2009.11.001

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology and Community Health*, *62*(8), 752–8. http://doi.org/10.1136/jech.2007.060798

Clason, D. L., Dormody, T. J., & Scales, L. (1993). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, *35*(4), 31–35.

Coe, R., & Fitz-Gibbon, C. T. (1998). School Effectiveness Research: Criticisms and Recommendations. *Oxford Review of Education*, *24*(4), 421–438. http://doi.org/10.2307/1050665

Coenders, M., & Scheepers, P. (2003). The Effect of Education on Nationalism and Ethnic Exclusionism: An International Comparison. *Political Psychology*, *24*(2), 313–343. http://doi.org/10.1111/0162-895X.00330

Cohen, J. (2012). School Climate and Culture Improvement. In P. M. Brown, M. W. Corrigan, & A. Higgins-D'Alessandro (Eds.), *Handbook of Prosocial Education* (pp. 227–270). Lanham, Maryland: Rowman & Littlefield Publishers, Inc.

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, Policy, Practice, and Teacher Education. *The Teachers College Record*, *111*(1), 180–213. Retrieved from https://www.tcrecord.org/Content.asp?ContentID=15220

Cohrs, J. C., Kielmann, S., Maes, J., & Moschner, B. (2005). Effects of right-wing authoritarianism and threat from terrorism on restriction of civil liberties. *Analyses of Social Issues and Public Policy*, *5*(1), 263–276. http://doi.org/10.1111/j.1530-2415.2005.00071.x

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity: Summary report* (Vol. 2). US Department of Health, Education, and Welfare, Office of Education.

Collado, D., Lomos, C., & Nicaise, I. (2014). The effects of classroom socioeconomic composition on student's civic knowledge in Chile. *School Effectiveness and School Improvement*, (January 2015), 1–26.

http://doi.org/10.1080/09243453.2014.966725

Collins, T. N., & Parson, K. A. (2010). School Climate and Student Outcomes. *Journal of Cross-Disciplinary Perspectives in Education*, *3*(1), 34–39.

Contreras, D., Elacqua, G., Martinez, M., & Miranda, Á. (2015). Income Inequality or Performance Gap? A Multilevel Study of School Violence in 52 Countries. *Journal of Adolescent Health*. http://doi.org/10.1016/j.jadohealth.2015.08.002

Cornell, D., Gregory, A., Huang, F., & Fan, X. (2013). Perceived prevalence of teasing and bullying predicts high school dropout rates. *Journal of Educational Psychology*, *105*(1), 138–149. http://doi.org/10.1037/a0030416

Cotton, J. L., & Tuttle, J. M. (1986). Employee Turnover: A Meta-Analysis and Review with Implications for Research. *Academy of Management Review*, *11*(1), 55–70. http://doi.org/10.5465/AMR.1986.4282625

Creemers, B., & Kyriakides, L. (2007). *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. London and New York: Routledge.

Creemers, B., & Kyriakides, L. (2010). School Factors Explaining Achievement on Cognitive and Affective Outcomes: Establishing a Dynamic Model of Educational Effectiveness. *Scandinavian Journal of Educational Research*, *54*(3), 263–294. http://doi.org/10.1080/00313831003764529

Cronbach, L. J. (1976). Research on Classrooms and Schools: Formulation of Questions, Design and Analysis. *Occasional Paper, Stanford Evaluation Consortium*. Retrieved from http://eric.ed.gov/?id=ED135801

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. http://doi.org/10.1037/h0040957

Cunningham, N. J. (2007). Level of Bonding to School and Perception of the School Environment by Bullies, Victims, and Bully Victims. *The Journal of Early Adolescence*, *27*(4), 457–478. http://doi.org/10.1177/0272431607302940

Desa, D. (2014). Evaluating Measurement Invariance of TALIS 2013 Complex Scales Comparison between Continuous and Categorical Multiple-Group Confirmatory Factor Analyses. OECD Education Working Papers. http://doi.org/10.1787/19939019

Dewey, J. (1966). *Democracy in Education*. *The Elementary School Teacher* (Vol. 4). New York: The Free Press.

Diekman, A. B., & Schneider, M. C. (2010). A Social Role Theory Perspective on Gender Gaps in Political Attitudes. *Psychology of Women Quarterly*, *34*(4), 486–497. http://doi.org/10.1111/j.1471-6402.2010.01598.x

Djonko-Moore, C. M. (2015). An exploration of teacher attrition and mobility in high poverty racially segregated schools. *Race Ethnicity and Education*, (August), 1–25. http://doi.org/10.1080/13613324.2015.1013458

Dolton, P., & Newson, D. (2003). The Relationship between Teacher Turnover and School Performance. *London Review of Education*, *1*(2), 132–140. http://doi.org/10.1080/14748460306685

Duckitt, J., Bizumic, B., Krauss, S. W., Heled, E., Duckitt, J., Bizumic, B., … Heled, E.

(2010). A Tripartite Approach to Right-Wing Authoritarianism : The Authoritarianism-Conservatism-Traditionalism. *Political Psychology*, *31*(5), 685–715.

Duckitt, J., & Sibley, C. G. (2009). A Dual-Process Motivational Model of Ideology, Politics, and Prejudice. *Psychological Inquiry*, *20*(2-3), 98–109. http://doi.org/10.1080/10478400903028540

Duckitt, J., Wagner, C., du Plessis, I., & Birum, I. (2002). The psychological bases of ideology and prejudice: Testing a dual process model. *Journal of Personality and Social Psychology*, *83*(1), 75–93. http://doi.org/10.1037//0022-3514.83.1.75

Due, P., Merlo, J., Harel-Fisch, Y., Damsgaard, M. T., Holstein, B. E., Hetland, J., … Lynch, J. (2009). Socioeconomic Inequality in Exposure to Bullying During Adolescence: A Comparative, Cross-Sectional, Multilevel Study in 35 Countries. *American Journal of Public Health*, *99*(5), 907–914. http://doi.org/10.2105/AJPH.2008.139303

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students’ Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, *82*(1), 405–432. http://doi.org/10.1111/j.1467-8624.2010.01564.x

Dweck, C. S., Chiu, C., & Hong, Y. (2009). Implicit Theories and Their Role in Judgments and Reactions: A Word From Two Perspectives. *Psychological Inquiry*, *6*(4), 267–285. http://doi.org/10.1207/s15327965pli0604_1

Dworkin, A. G. (1987). *Teacher Burnout in the Public Schools: Structural Causes and Consequences for Children*. New York, NY: SUNY Press.

Dworkin, A. G. (2001). Perspectives on teacher burnout and school reform. *International Education Journal*, *2*(2), 69–78.

Easterbrook, M. J., Kuppens, T., & Manstead, A. S. R. (2015). The Education Effect: Higher Educational Qualifications are Robustly Associated with Beneficial Personal and Socio-political Outcomes. *Social Indicators Research*. http://doi.org/10.1007/s11205-015-0946-1

Edmonds, R. R. (1979a). Effective schools for the urban poor. *Educational Leadership*, *37*(1), 15–18.

Edmonds, R. R. (1979b). Some schools work and more can. *Social Policy*, *9*(5), 28–32.

Edmonds, R. R. (1982). *Programs of school improvement: An overview*. Washington, DC.: National Inst. of Education (ED). Retrieved from http://files.eric.ed.gov/fulltext/ED221536.pdf

Eggleston, J. (1977). *The Ecology of the School*. New York, NY: Routledge.

Elacqua, G. (2012). The impact of school choice and public policy on segregation: Evidence from Chile. *International Journal of Educational Development*, *32*(3), 444–453. http://doi.org/10.1016/j.ijedudev.2011.08.003

Elgar, F. J., Craig, W., Boyce, W., Morgan, A., & Vella-Zarb, R. (2009). Income Inequality and School Bullying: Multilevel Study of Adolescents in 37 Countries. *Journal of Adolescent Health*, *45*(4), 351–359. http://doi.org/10.1016/j.jadohealth.2009.04.004

Enders, C. K. (2011). Missing not at random models for latent growth curve analyses. *Psychological Methods*, *16*(1), 1–16. http://doi.org/10.1037/a0022640

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, *12*(2), 121–38. http://doi.org/10.1037/1082-989X.12.2.121

Engel, L. C., Rutkowski, D., & Rutkowski, L. (2009). The harsher side of globalisation: violent conflict and academic achievement. *Globalisation, Societies and Education*, *7*(4), 433–456. http://doi.org/10.1080/14767720903412242

Espelage, D. L., & Holt, M. K. (2013). Suicidal Ideation and School Bullying Experiences After Controlling for Depression and Delinquency. *Journal of Adolescent Health*, *53*(1), S27–S31. http://doi.org/10.1016/j.jadohealth.2012.09.017

Espelage, D. L., Holt, M. K., & Henkel, R. R. (2003). Examination of Peer-Group Contextual Effects on Aggression During Early Adolescence. *Child Development*, *74*(1), 205–220. http://doi.org/10.1111/1467-8624.00531

Espelage, D. L., Hong, J. S., Rao, M. A., & Low, S. (2013). Associations Between Peer Victimization and Academic Performance. *Theory Into Practice*, *52*(4), 233–240. http://doi.org/10.1080/00405841.2013.829724

Evans, G., & Rose, P. (2007). Support for Democracy in Malawi: Does Schooling Matter? *World Development*, *35*(5), 904–919. http://doi.org/10.1016/j.worlddev.2006.09.011

Evans, G., & Rose, P. (2012). Understanding Education's Influence on Support for Democracy in Sub-Saharan Africa. *Journal of Development Studies*, *48*(4), 498–515. http://doi.org/10.1080/00220388.2011.598508

Eyzaguirre, S. (2015, August 3). Cómo atraer a docentes talentosos a trabajar en escuelas vulnerables. *El Mostrador*. Santiago, Chile. Retrieved from http://www.elmostrador.cl/noticias/opinion/2015/08/03/como-atraer-a-docentes-talentosos-a-trabajar-en-escuelas-vulnerables/

Farrington, D. P., & Ttofi, M. M. (2011). Bullying as a predictor of offending, violence and later life outcomes. *Criminal Behaviour and Mental Health*, *21*(2), 90–98. http://doi.org/10.1002/cbm.801

Feaster, D., Brincks, A., Robbins, M., & Szapocznik, J. (2011). Multilevel Models to Identify Contextual Effects on Individual Group Member Outcomes: A Family Example. *Family Process*, *50*(2), 167–183. http://doi.org/10.1111/j.1545-5300.2011.01353.x

Feldman, S. (2003). Enforcing Social Conformity: A Theory of Authoritarianism. *Political Psychology*, *24*(1), 41–74. http://doi.org/10.1111/0162-895X.00316

Feng, L. (2010). Hire Today, Gone Tomorrow: New Teacher Classroom Assignments and Teacher Mobility. *Education Finance and Policy*, *5*(3), 278–316. http://doi.org/10.1162/EDFP_a_00002

Finster, M. P. (2013). *Teachers' Job Satisfaction, Organizational Commitment, Turnover Intentions, and Actual Turnover: A Secondary Analysis using an Integrative Structural Equation Modeling Approach*. University of Washington. Retrieved from http://hdl.handle.net/1773/23621

Fleming, L. C., & Jacobsen, K. H. (2009). Bullying and symptoms of depression in chilean middle school students. *Journal of School Health*, *79*(3), 130–137. http://doi.org/10.1111/j.1746-1561.2008.0397.x

Forsyth, P. B., Adams, C. M., & Hoy, W. K. (2011). *Collective Trust: Why Schools Can't Improve Without it*. New York, NY: Teachers College Press.

Foy, P., Brossman, B., & Galia, J. (2011). Scaling the TIMSS and PIRLS 2011 achievement data. *TIMSS and PIRLS Methods and Procedures*, 1–28. Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP11_Scaling_Achievement.pdf

Fraillon, J., Schulz, W., & Ainley, J. (2012). *ICCS 2009 Asian Report*. International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Asian_Report.pdf

Fraser, B. J. (1989). Twenty years of classroom climate work: progress and prospect. *Journal of Curriculum Studies*, *21*(4), 307–327. http://doi.org/10.1080/0022027890210402

Freiberg, H. J. (1998). Measuring school climate: Let me count the ways. *Educational Leadership*, *56*(1), 22.

Freiberg, H. J. (1999). *School climate: Measuring, improving, and sustaining healthy learning environments*. London, UK: Falmer Press.

Funke, F. (2005). The dimensionality of right-wing authoritarianism: Lessons from the dilemma between theory and measurement. *Political Psychology*, *26*(2), 195–218. http://doi.org/10.1111/j.1467-9221.2005.00415.x

Geiser, C. (2012). *Data Analysis with Mplus*. New York and London: The Guilford Press.

Gelman, A. (2006). Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. *Technometrics*, *48*(3), 432–435. http://doi.org/10.1198/004017005000000661

Gelman, A. (2007). Clustered standard errors vs. multilevel modeling. Retrieved from http://andrewgelman.com/2007/11/28/clustered_stand/

Gill, M. G., Ashton, P., & Algina, J. (2004). Authoritative schools: A test of a model to resolve the school effectiveness debate. *Contemporary Educational Psychology*, *29*(4), 389–409. http://doi.org/10.1016/j.cedpsych.2003.10.002

Glover, D., & Coleman, M. (2005). School culture, climate and ethos: interchangeable or distinctive concepts? *Journal of In-Service Education*, *31*(2), 251–272. http://doi.org/10.1080/13674580500200278

Goddard, R., & O'Brien, P. (2003). Beginning teacher perceptions of their work, well-being and intention to leave. *Asia-Pacific Journal of Teacher Education and Development*, *6*(2), 99–110. Retrieved from http://libdr1.ied.edu.hk/pubdata/img00/arch00/link/archive/1/instarh/3015_image_vol6_no2_p99_168.pdf

Godfrey, E. B., & Grayman, J. K. (2014). Teaching Citizens: The Role of Open Classroom Climate in Fostering Critical Consciousness Among Youth. *Journal of Youth and Adolescence*, *43*(11), 1801–1817. http://doi.org/10.1007/s10964-013-0084-5

Goldhaber, D., & Hansen, M. (2009). National Board Certification and Teachers' Career Paths: Does NBPTS Certification Influence How Long Teachers Remain in the Profession and Where They Teach? *Education Finance and Policy*, *4*(3), 229–262. http://doi.org/10.1162/edfp.2009.4.3.229

Goldstein, H., & Woodhouse, G. (2000). School Effectiveness Research and Educational Policy. *Oxford Review of Education*, *26*(3-4), 353–363. http://doi.org/10.1080/713688547

Goodlad, J. I., & Sava, S. G. (1975). *The dynamics of educational change: Toward responsive schools*. New York, NY: McGraw-Hill.

Gorfein, D. (1961). Conformity Behavior and the "Authoritarian Personality." *The Journal of Social Psychology*, *53*(1), 121–125. http://doi.org/10.1080/00224545.1961.9922108

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator Tests, and Research Implications for the Next Millennium. *Journal of Management*, *26*(3), 463–488. http://doi.org/10.1177/014920630002600305

Griffith, J. (1995). An Empirical Examination of a Model of Social Climate in Elementary Schools. *Basic and Applied Social Psychology*, *17*(1-2), 97–117. http://doi.org/10.1080/01973533.1995.9646134

Griffith, J. (1997). School Climate as "Social Order" and "Social Action": A Multi-Level Analysis of Public Elementary School Student Perceptions. *Social Psychology of Education*, *2*(3-4), 339–369. http://doi.org/10.1023/A:1009657422344

Griffith, J. (2000). School Climate as Group Evaluation and Group Consensus: Student and Parent Perceptions of the Elementary School Environment. *The Elementary School Journal*, *101*(1), 35–61. http://doi.org/10.1086/499658

Grissmer, D., & Kirby, S. (1992). *Patterns of attrition among Indiana teachers, 1965-1987*. Santa Monica, CA: RAND. Retrieved from https://www.rand.org/content/dam/rand/pubs/reports/2007/R4076.pdf

Grissmer, D., & Kirby, S. (1997). Teacher turnover and teacher quality. *The Teachers College Record*, *99*(1), 45–56.

Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher Recruitment and Retention: A Review of the Recent Empirical Literature. *Review of Educational Research*, *76*(2), 173–208. http://doi.org/10.3102/00346543076002173

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136. http://doi.org/10.1037/0022-0663.95.1.124

Guimond, S., Dambrun, M., Michinov, N., & Duarte, S. (2003). Does social dominance generate prejudice? Integrating individual and contextual determinants of intergroup cognitions. *Journal of Personality and Social Psychology*, *84*(4), 697–721. http://doi.org/10.1037/0022-3514.84.4.697

Guin, K. (2004). Chronic Teacher Turnover in Urban Elementary Schools. *Education Policy Analysis Archives*, *12*, 42. http://doi.org/10.14507/epaa.v12n42.2004

Gutman, L. M., & Midgley, C. (2000). The Role of Protective Factors in Supporting the Academic Achievement of Poor African American Students During the Middle School Transition. *Journal of Youth and Adolescence*, *29*(2), 223–249. http://doi.org/10.1023/A:1005108700243

Hakanen, J. J., Bakker, A. B., & Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of School Psychology*, *43*(6), 495–513. http://doi.org/10.1016/j.jsp.2005.11.001

Hallquist, M., & Wiley, J. (2012). Mplusautomation: Automating mplus model estimation and interpretation. Computer software manual. Retrieved from http://cran.r-project.org/web/packages/MplusAutomation/index.html

Halpin, A. W., & Croft, D. B. (1963). *The organizational climate of schools*. Midewest Administration Center, University of Chicago.

Handy, C. B., & Aitken, R. (1986). *Understanding schools as organizations*. London: Penguin books.

Hanh, C. (1998). *Becoming Political: Comparative Perspectives on Citizenship Education*. Albany, NY, United States: SUNY Press.

Hanisch, K. A., & Hulin, C. L. (1991). General attitudes and organizational withdrawal: An evaluation of a causal model. *Journal of Vocational Behavior*, *39*(1), 110–128. http://doi.org/10.1016/0001-8791(91)90006-8

Harrison, D. A., Newman, D. A., & Roth, P. L. (2006). How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal*, *49*(2), 305–325. http://doi.org/10.5465/AMJ.2006.20786077

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press.

Hayes, A. F. (2015). An Index and Simple Test of Linear Moderated Mediation. *Multivariate Behavioral Research*, *50*(1), 1–54. http://doi.org/10.1080/00273171.2014.962683

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches Using MPLUS*. New York and Hove: Routledge.

Heeringa, S. G., West, B., & Berglund, P. A. (2009). *Applied Survey Data Analysis*. Boca Raton, London, New York: Taylor & Francis Group.

Hertz, M. F., Donato, I., & Wright, J. (2013). Bullying and Suicide: A Public Health Approach. *Journal of Adolescent Health*, *53*(1), S1–S3. http://doi.org/10.1016/j.jadohealth.2013.05.002

Hibbing, J. R., & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Chicago, Illinois: Cambridge University Press.

Highton, B. (2009). Revisiting the Relationship between Educational Attainment and Political Sophistication. *The Journal of Politics*, *71*(04), 1564. http://doi.org/10.1017/S0022381609990077

Hirsch, E., Koppich, J. E., & Knapp, M. S. (2001). *Revisiting What States are Doing to Improve the Quality of Teaching : An Update on Patterns and Trends* (Vol. 2001).

Seattle, WA. Retrieved from http://depts.washington.edu/ctpmail/PDFs/States-HKK-02-2001.pdf

Hom, P. W., Caranikas-Walker, F., Prussia, G. E., & Griffeth, R. W. (1992). A meta-analytical structural equations analysis of a model of employee turnover. *Journal of Applied Psychology*, *77*(6), 890–909. http://doi.org/10.1037/0021-9010.77.6.890

Hooghe, M., Meeusen, C., & Quintelier, E. (2013). The impact of education and intergroup friendship on the development of ethnocentrism. A latent growth curve model analysis of a five-year panel study among belgian late adolescents. *European Sociological Review*, *29*(6), 1109–1121. http://doi.org/10.1093/esr/jcs086

Horng, E. L. (2009). Teacher Tradeoffs: Disentangling Teachers' Preferences for Working Conditions and Student Demographics. *American Educational Research Journal*, *46*(3), 690–717. http://doi.org/10.3102/0002831208329599

Houtman, D. (2003). Lipset and "working-class" authoritarianism. *The American Sociologist*, *34*(1-2), 85–103. http://doi.org/10.1007/s12108-003-1008-8

Houtte, M. Van, & Stevens, P. A. J. (2008). Sense of Futility The Missing Link Between Track Position and Self-Reported School Misconduct. *Youth & Society*, *40*(2), 245–264. http://doi.org/10.1177/0044118X08316251

Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications*. New York and Hove: Routledge.

Hoy, W. K. (2012). School characteristics that make a difference for the achievement of all students: A 40-year odyssey. *Journal of Educational Administration*, *50*(1), 76–97. http://doi.org/10.1108/09578231211196078

Hoy, W. K., & Fedman, J. A. (1987). Organizational Health: The Concept and Its Measure. *Journal of Research and Development in Education*, *20*(4), 30–37.

Huck, S. W. (2012). *Reading Statistics and Research* (6th ed.). Boston, MA: Pearson Education.

Ichilov, O. (1991). Political Socialization and Schooling Effects among Israeli Adolescents. *Comparative Education Review*, *35*(3), 430. http://doi.org/10.1086/447046

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. http://doi.org/10.1037/a0020761

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, *25*(1), 51–71. http://doi.org/10.1214/10-STS321

Imai, K., King, G., & Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *171*(2), 481–502. http://doi.org/10.1111/j.1467-985X.2007.00527.x

Imazeki, J. (2005). Teacher salaries and teacher attrition. *Economics of Education Review*, *24*(4), 431–449. http://doi.org/10.1016/j.econedurev.2004.07.014

Ingersoll, R. M. (2001). Teacher Turnover and Teacher Shortages: An Organizational Analysis. *American Educational Research Journal*, *38*(3), 499–534.

http://doi.org/10.3102/00028312038003499

Ingersoll, R. M. (2002). The Teacher Shortage: A Case of Wrong Diagnosis and Wrong Prescription. *NASSP Bulletin*, *86*(631), 16–31. http://doi.org/10.1177/019263650208663103

Ingersoll, R. M., & May, H. (2012). The Magnitude, Destinations, and Determinants of Mathematics and Science Teacher Turnover. *Educational Evaluation and Policy Analysis*, *34*(4), 435–464. http://doi.org/10.3102/0162373712454326

Ingersoll, R. M., & Smith, T. (2003). The wrong solution to the teacher shortage. *Educational Leadership*, *60*(8), 30–33.

INJUV. (2013). *7ma Encuesta Nacional de Juventud 2012*. Santiago, Chile: Ministerio de Desarrollo Social, Gobierno de Chile. Retrieved from http://www.injuv.gob.cl/portal/wp-content/files_mf/septimaencuestanacionaljuventud2.pdf

Isaacs, J., Voeten, M., & Salmivalli, C. (2013). Gender-specific or Common Classroom Norms? Examining the Contextual Moderators of the Risk for Victimization. *Social Development*, *22*(3), 555–579. http://doi.org/10.1111/j.1467-9507.2012.00655.x

Jackman, R. W. (1985). Cross-National Statistical Research and the Study of Comparative Politics. *American Journal of Political Science*, *29*(1), 161–182. http://doi.org/10.2307/2111217

Jansen, P. W., Verlinden, M., Dommisse-van Berkel, A., Mieloo, C., van der Ende, J., Veenstra, R., … Tiemeier, H. (2012). Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter? *BMC Public Health*, *12*(1), 494. http://doi.org/10.1186/1471-2458-12-494

Jencks, C. (1975). *Inequality: A reassessment of the effect of family and schooling in America.* London, United Kingdom: Penguin Books Ltd.

Jensen, B., Sandoval-Hernández, A., Knoll, S., & Gonzalez, E. J. (2012). *The Experience of New Teachers*. *OECD Publishing*. OECD Publishing. http://doi.org/10.1787/9789264120952-en

Johnson, D., Terry, D. J., & Louis, W. R. (2005). Perceptions of the Intergroup Structure and Anti-Asian Prejudice Among White Australians. *Group Processes & Intergroup Relations*, *8*(1), 53–71. http://doi.org/10.1177/1368430205048616

Joncas, M., & Foy, P. (2011). Sample Design in TIMSS and PIRLS. *TIMSS and PIRLS Methods and Procedures*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo. *Political Psychology*, *25*(6), 881–919. http://doi.org/10.1111/j.1467-9221.2004.00402.x

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*, 307–37.

http://doi.org/10.1146/annurev.psych.60.110707.163600

JUNAEB. (2005). *SINAE Sistema Nacional de Asignación con Equidad para Becas JUNAEB. Una nueva visión en la construcción de igualdad de oportunidades en la infancia*. Santiago, Chile: Gobierno de Chile, JUNAEB. Retrieved from http://www.junaeb.cl/wp-content/uploads/2013/02/libro_junaeb.pdf

Kain, J., Uauy, R., & Taibo, M. (2002). Chile's school feeding programme: Targeting experience. *Nutrition Research*, *22*(5), 599–608. http://doi.org/10.1016/S0271-5317(01)00377-3

Kammeyer-Mueller, J. D., & Wanberg, C. R. (2003). Unwrapping the organizational entry process: disentangling multiple antecedents and their pathways to adjustment. *The Journal of Applied Psychology*, *88*(5), 779–94. http://doi.org/10.1037/0021-9010.88.5.779

Kammeyer-Mueller, J. D., Wanberg, C. R., Glomb, T. M., & Ahlburg, D. (2005). The role of temporal shifts in turnover processes: it's about time. *The Journal of Applied Psychology*, *90*(4), 644–58. http://doi.org/10.1037/0021-9010.90.4.644

Kaplan, D. (2014). Bayesian Statistics for the Social Sciences. New York, NY: Guilford Press.

Kelley, E., Glover, J., Keefe, J., Halderson, C., Sorenson, C., & Speth, C. (1986). School Climate Survey (Modified) Form A: Examiner's manual. *Reston, VA: NAASP*.

Kelly, S. (2004). An Event History Analysis of Teacher Attrition: Salary, Teacher Tracking, and Socially Disadvantaged Schools. *The Journal of Experimental Education*, *72*(3), 195–220. http://doi.org/10.3200/JEXE.72.3.195-220

Kim, J. S., Anderson, C. J., & Keller, B. (2013). Multilevel Analysis of Assessment Data. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analaysis* (pp. 389–424). Boca Raton, London, New York: Chapman and Hall/CRC.

Kirby, S., Naftel, S., & Berends, M. (1999). *Staffing at-risk school districts in Texas: Problems and prospects*. Santa Monica, CA: RAND. Retrieved from http://www.rand.org/content/dam/rand/pubs/monograph_reports/2007/MR1083.pdf

Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling* (3rd ed.). London: The Guilford Press.

Konishi, C., Hymel, S., Zumbo, B. D., & Li, Z. (2010). Do school bullying and student-teacher relationships matter for academic achievement? A Multilevel Analysis. *Canadian Journal of School Psychology*, *25*(1), 19–39. http://doi.org/10.1177/0829573509357550

Konu, A. I., Lintonen, T. P., & Autio, V. J. (2002). Evaluation of Well-Being in Schools? A Multilevel Analysis of General Subjective Well-Being. *School Effectiveness and School Improvement*, *13*(2), 187–200. http://doi.org/10.1076/sesi.13.2.187.3432

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology*, *100*(1), 96–104. http://doi.org/10.1037/0022-

0663.100.1.96

Kruglanski, A. W., Pierro, A., Mannetti, L., & De Grada, E. (2006). Groups as epistemic providers: need for closure and the unfolding of group-centrism. *Psychological Review*, *113*(1), 84–100. http://doi.org/10.1037/0033-295X.113.1.84

Kukla-Acevedo, S. (2009). Leavers, Movers, and Stayers: The Role of Workplace Conditions in Teacher Mobility Decisions. *The Journal of Educational Research*, *102*(6), 443–452. http://doi.org/10.3200/JOER.102.6.443-452

Kyriakides, L., & Creemers, B. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: a study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, *19*(2), 183–205. http://doi.org/10.1080/09243450802047873

Kyriakides, L., Creemers, B., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal*, *36*(5), 807–830. http://doi.org/10.1080/01411920903165603

Lacey, A., & Cornell, D. (2013). The Impact of Teasing and Bullying on Schoolwide Academic Performance. *Journal of Applied School Psychology*, *29*(3), 262–283. http://doi.org/10.1080/15377903.2013.806883

Ladd, H. F. (2011). Teachers' Perceptions of Their Working Conditions: How Predictive of Planned and Actual Teacher Movement? *Educational Evaluation and Policy Analysis*, *33*(2), 235–261. http://doi.org/10.3102/0162373711398128

Lamote, C., Van Damme, J., Van Den Noortgate, W., Speybroeck, S., Boonen, T., & de Bilde, J. (2013). Dropout in secondary education: an application of a multilevel discrete-time hazard model accounting for school changes. *Quality & Quantity*, *47*(5), 2425–2446. http://doi.org/10.1007/s11135-012-9662-y

Lancee, B., & Sarrasin, O. (2015). Educated Preferences or Selection Effects? A Longitudinal Analysis of the Impact of Educational Attainment on Attitudes Towards Immigrants. *European Sociological Review*, *31*(4), 490–501. http://doi.org/10.1093/esr/jcv008

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62. http://doi.org/10.3102/01623737024001037

Lauder, H., Jamieson, I., & Wikeley, F. (1998). Models of Effective Schools: Limits and Capabilities. In R. Slee, G. Weiner, & S. Tomlinson (Eds.), *School Effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements*. London: Falmer Press.

Lauglo, J. (2013). Do more knowledgeable adolescents have more rationally based civic attitudes? Analysis of 38 countries. *Educational Psychology*, *33*(3), 262–282. http://doi.org/10.1080/01443410.2013.772773

Lee, J.-S. (2014). The Relationship Between Student Engagement and Academic Performance: Is It a Myth or Reality? *The Journal of Educational Research*, *107*(3), 177–185. http://doi.org/10.1080/00220671.2013.807491

Lee, T. H. (2013). Temporal aspects of job mobility: A comparison of traditional and survival analysis. *Journal of Management & Marketing Research*, *12*, 1–12.

Lee, V. E., & Bryk, A. S. (1988). Curriculum Tracking as Mediating the Social Distribution of High School Achievement. *Sociology of Education*, *61*(2), 78–94. http://doi.org/10.2307/2112266

Lenzi, M., Vieno, A., Sharkey, J., Pastore, M., & Santinello, M. (2014). How School can Teach Civic Engagement Besides Civic Education : The Role of Democratic School Climate. *Am J Community Psychol*, (54), 251–261. http://doi.org/10.1007/s10464-014-9669-8

Lewin, K., Heider, F. T., & Heider, G. M. (1936). *Principles of topological psychology*. New York and London: McGraw-Hill Book Company.

Lewin, K., & Lippitt, R. (1938). An Experimental Approach to the Study of Autocracy and Democracy: A Preliminary Note. *Sociometry*, *1*(3/4), 292. http://doi.org/10.2307/2785585

Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of Aggressive Behavior in Experimentally Created “Social Climates.” *The Journal of Social Psychology*, *10*(2), 269–299. http://doi.org/10.1080/00224545.1939.9713366

Li, C.-H. (2015). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*. http://doi.org/10.3758/s13428-015-0619-7

Lindahl, R. (2006). *The role of organizational climate and culture in the school improvement process: A review of the knowledge base*. National Council of Professors of Educational Administration. Retrieved from cnx.org/content/m13465/1.1/

Lippitt, R., & White, R. K. (1943). The “social climate” of children’s groups. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds.), *Child behavior and development: A course of representative studies.* (pp. 485–508). New York, NY, US: McGraw-Hill.

Lipset, S. M. (1959). Democracy and Working-Class Authoritarianism. *American Sociological Review*, *24*(4), 482–501. Retrieved from http://www.jstor.org/stable/2089536

Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: model specification and estimation. *British Educational Research Journal*, *41*(3), 423–447. http://doi.org/10.1002/berj.3147

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How Teaching Conditions Predict Teacher Turnover in California Schools. *Peabody Journal of Education*, *80*(3), 44–70. http://doi.org/10.1207/s15327930pje8003_4

Lucas, S. R. (2014). An inconvenient dataset: bias and inappropriate inference with the multilevel model. *Quality & Quantity*, *48*, 1619–1649. http://doi.org/10.1007/s11135-013-9865-x

Luchman, J. N. (2014). *Cluster-level Correlated Error Variance and the Estimation of Parameters in Linear Mixed Models*. George Mason University. Retrieved from http://hdl.handle.net/1920/8835

Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 × 2 taxonomy of multilevel latent contextual models: accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, *16*(4), 444–67.

http://doi.org/10.1037/a0024376

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203–229. http://doi.org/10.1037/a0012869

Luskin, R. C. (1990). Explaining political sophistication. *Political Behavior*, *12*(4), 331–361. http://doi.org/10.1007/BF00992793

Luyten, H., Visscher, A., & Witziers, B. (2005). School Effectiveness Research: From a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement*, *16*(3), 249–279. http://doi.org/10.1080/09243450500114884

Ma, X., Ma, L., & Bradley, K. D. (2008). Using multilevel modeling to investigate school effects. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data educational data* (pp. 59–110). Charlotte, NC: IAP.

Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, *58*(2), 127–137. http://doi.org/10.1046/j.0039-0402.2003.00252.x

Maas, C. J. M., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *1*(3), 86–92. http://doi.org/10.1027/1614-2241.1.3.86

MacAulay, D. J. (1990). Classroom Environment: a literature review. *Educational Psychology*, *10*(3), 239–253. http://doi.org/10.1080/0144341900100305

Macdonald, D. (1999). Teacher attrition: a review of literature. *Teaching and Teacher Education*, *15*(8), 835–848. http://doi.org/10.1016/S0742-051X(99)00031-1

Macdonald, K. (2014). PV: Stata module to perform estimation with plausible values. *Statistical Software Components*.

Madriaza, P. (2008). Violencia escolar en Chile. In C. Guajardo (Ed.), *Seguridad y Prevención: La situación en Argentina, Chile y Uruguay durante 2007* (pp. 114–139). Retrieved from http://ww.jhc.cl/derecho/wp-content/uploads/2013/06/Seguridad_y_Prevencion.pdf#page=116

Magendzo, A., Toledo, M., & Gutiérrez, V. (2012). Descripción y análisis de la Ley sobre Violencia Escolar (Nº20 . 536): dos paradigmas antagónicos. *Estudios Pedagógicos (Valdivia)*, *39*(1), 377–391. Retrieved from http://www.scielo.cl/pdf/estped/v39n1/art22.pdf

Manzi, J., Gutiérrez, R. G., & Sun, Y. (2011). *La evaluación docente en Chile*. Santiago, Chile: MIDE UC.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.

Markus, K. A. (2008). Constructs, Concepts and the Worlds of Possibility: Connecting the Measurement, Manipulation, and Meaning of Variables. *Measurement: Interdisciplinary Research & Perspective*, *6*(1-2), 54–77. http://doi.org/10.1080/15366360802035513

Marsh, H. W., Abduljabbar, a. S., Parker, P. D., Morin, a. J. S., Abdelfattah, F., & Nagengast, B. (2014). The Big-Fish-Little-Pond Effect in Mathematics: A Cross-

Cultural Comparison of U.S. and Saudi Arabian TIMSS Responses. *Journal of Cross-Cultural Psychology*, *45*(5), 777–804. http://doi.org/10.1177/0022022113519858

Marsh, H. W., Kuyper, H., Morin, A. J. S., Parker, P. D., & Seaton, M. (2014). Big-fish-little-pond social comparison and local dominance effects: Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction*, *33*, 50–66. http://doi.org/10.1016/j.learninstruc.2014.04.002

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom Climate and Contextual Effects: Conceptual and Methodological Issues in the Evaluation of Group-Level Effects. *Educational Psychologist*, *47*(2), 106–124. http://doi.org/10.1080/00461520.2012.670488

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-Latent Models of School Contextual Effects: Integrating Multilevel and Structural Equation Approaches to Control Measurement and Sampling Error. *Multivariate Behavioral Research*, *44*(6), 764–802. http://doi.org/10.1080/00273170903333665

Marsh, H. W., Trautwein, U., Ludtke, O., Baumert, J., & Koller, O. (2007). The Big-Fish-Little-Pond Effect: Persistent Negative Effects of Selective High Schools on Self-Concept After Graduation. *American Educational Research Journal*, *44*(3), 631–669. http://doi.org/10.3102/0002831207306728

Martin, M. O., Mullis, I. V. S., Foy, P., & Arora, A. (2011). Creating and Interpreting the TIMSS and PIRLS 2011 Context Questionnaire Scales. In *TIMSS 2011 Assessment Frameworks* (pp. 1–11). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP11_Context_Q_Scales.pdf

Maslow, A. H. (2002). *Psychology of Science : A Reconnaissance*. *The John Dewey Society lectureship series*. Maurice Bassett.

Maslowski, R. (2006). A review of inventories for diagnosing school culture. *Journal of Educational Administration*, *44*(1), 6–35. http://doi.org/10.1108/09578230610642638

Masyn, K. E. (2003). *Discrete-Time Survival Mixture Analysis for Single and Recurrent Events Using Latent Variables*. University of California Los Angeles. Retrieved from http://www.statmodel.com/download/masyndissertation.pdf

Masyn, K. E. (2014). Discrete-Time Survival Analysis In prevention Science. In Z. Sloboda & H. Petras (Eds.), *Defining Prevention Science* (pp. 513–535). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4899-7424-2

Mavor, K. I., Louis, W. R., & Sibley, C. G. (2010). A bias-corrected exploratory and confirmatory factor analysis of right-wing authoritarianism: Support for a three-factor structure. *Personality and Individual Differences*, *48*(1), 28–33. http://doi.org/10.1016/j.paid.2009.08.006

McEvoy, A., & Welker, R. (2000). Antisocial Behavior, Academic Failure, and School Climate: A Critical Review. *Journal of Emotional and Behavioral Disorders*, *8*(3), 130–140. http://doi.org/10.1177/106342660000800301

McEwan, P. J. (2010). *The impact of school meals on education outcomes: discontinuity evidence from Chile*. *Department of Economics, Wellesley College, MA*. Department of Economics, Wellesley College, MA. Retrieved from https://aefpweb.org/sites/default/files/webform/mcewan.pdf

McEwan, P. J., Urquiola, M., & Vegas, E. (2008). School Choice, Stratification, and Information on School Performance: Lessons from Chile. *Economía*, *8*(2), 1–27. http://doi.org/10.1353/eco.0.0003

McFarland, S. (2010). Personality and support for universal human rights: a review and test of a structural model. *Journal of Personality*, *78*(6), 1735–63. http://doi.org/10.1111/j.1467-6494.2010.00668.x

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology*, *4*(1), 103–120. http://doi.org/10.1080/0022250X.1975.9989847

McNeish, D. M. (2014). Modeling sparsely clustered data: Design-based, model-based, and single-level methods. *Psychological Methods*, *19*(4), 552–563. http://doi.org/10.1037/met0000024

McNeish, D. M., & Harring, J. R. (2015). Clustered Data with Small Sample Sizes: Comparing the Performance of Model-based and Design-based Approaches. *Communications in Statistics - Simulation and Computation*. http://doi.org/10.1080/03610918.2014.983648

McNeish, D. M., & Stapleton, L. M. (2014). The Effect of Small Sample Size on Two-Level Model Estimates: A Review and Illustration. *Educational Psychology Review*. http://doi.org/10.1007/s10648-014-9287-x

McNeish, D. M., & Stapleton, L. M. (2015). Clustered data mean you need multilevel models, right? In *Annual Meeting of the American Educational Research Association (AERA) SIG: Multilevel Modeling.* Chicago, IL,USA.

Meckes, L., & Bascopé, M. (2012). Uneven Distribution of Novice Teachers in the Chilean Primary School System. *Education Policy Analysis Archives*, *20*(30), 30. http://doi.org/10.14507/epaa.v20n30.2012

Medley, D. M., & Mitzel, H. E. (1958). Application of Analysis of Variance to the Estimation of the Reliability of Observations of Teachers' Classroom Behavior. *The Journal of Experimental Education*, *27*(1), 23–35. http://doi.org/10.2307/20154100

Medley, D. M., & Mitzel, H. E. (1963). Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 247–328). Chicago: Rand McNally.

Mehta, S. B., Cornell, D., Fan, X., & Gregory, A. (2013). Bullying climate and school engagement in ninth-grade students. *The Journal of School Health*, *83*(1), 45–52. http://doi.org/10.1111/j.1746-1561.2012.00746.x

Meuleman, B., & Billiet, J. (2009). A Monte Carlo sample size study: How many countries are needed for accurate multilevel SEM? *Survey Research Methods*, *3*(1), 45–58. http://doi.org/10.18148/srm/2009.v3i1.666

Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Human Resource Management Review*, *1*(1), 61–89.

http://doi.org/10.1016/1053-4822(91)90011-Z

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., … Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182. http://doi.org/10.1126/science.1199644

Miller, S. I., & Fredericks, J. (1990). The False Ontology of School Climate Effects. *Educational Theory*, *40*(3), 333–342. http://doi.org/10.1111/j.1741-5446.1990.00333.x

Miller, S. M. (1973). On the Uses, Misuses and Abuses of Jencks' Inequality. *Sociology of Education*, *46*(4), 427–432. http://doi.org/10.2307/2111897

MINEDUC. (2004a). *Formación ciudadana en el currículum de la reforma*. Ministerio de Educación, República de Chile.

MINEDUC. (2004b). *Informe comisión formación ciudadana*. Santiago, Chile: Ministerio de Educación, Cultura y Deporte, Gobierno de España. Retrieved from http://wwwfs.mineduc.cl/Archivos/ConvivenciaEscolar/doc/archivo_153.pdf

MINEDUC. (2011a). *Ley 20.536 Sobre violencia escolar*. (MINEDUC, Ed.). Santiago, Chile: Biblioteca del Congreso Nacional de Chile. Retrieved from http://www.leychile.cl/Navegar?idNorma=1030087&idVersion=2011-09-17

MINEDUC. (2012). *Plan Escuela Segura*. (MINEDUC, Ed.). Santiago, Chile: Ministerio de Educación, República de Chile. Retrieved from http://www.comunidadescolar.cl/boletines_comunidad/2012/agosto/PDF 1 Escuela Segura.pdf

MINEDUC. (2014). *Informe Final: Programa Beca Vocación de Profesor*. Santiago de Chile. Retrieved from http://geminis.dipres.cl/virlib/docs/Gestion/2014/doc21-217-2014927_08_Beca_Vocacion_Profesor_info_fin.pdf

MINEDUC. (2015). Nueva Carrera Docente: Sistema de promoción y desarrollo professional docente. Retrieved from http://www.oei.es/formaciondocente/legislacion/CHILE/OTROS/PROYECTO_LEY_PROFESORADO.pdf

MINEDUC, C. de E. (2011b). *Estadísticas de la Educación 2011*. Santigo, Chile: Ministerio de Educación, República de Chile.

Mitchell, T. R., & Lee, T. W. (2001). The unfolding model of voluntary turnover and job embeddedness: Foundations for a comprehensive theory of attachment. *Research in Organizational Behavior*, *23*, 189–246. http://doi.org/10.1016/S0191-3085(01)23006-8

Mizala, A., & Romaguera, P. (2000). Remuneraciones al pizarrón. *Perspectivas En Política, Economía Y Gestión*, *4*(1), 65–88. Retrieved from http://www.dii.uchile.cl/~Revista/ArticulosVol4-N1/Mizala-A y Romaguera-P.pdf

Mizala, A., & Torche, F. (2012). Bringing the schools back in: The stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, *32*(1), 132–144. http://doi.org/10.1016/j.ijedudev.2010.09.004

Mo, L., Yang, F., & Hu, X. (2011). An empirical examination of IRT information for school climate surveys. *Educational Research and Evaluation*, *17*(1), 33–45.

http://doi.org/10.1080/13803611.2011.583033

Moos, R. H. (1979). *Evaluating educational environments: Measures, procedures, findings, and policy implications*. San Francisco, CA: Jossey-Bass Publishers.

Moos, R. H. (1980). *The social climate scales: An overview*. Palo Alto, California: Consulting Psychologists Press.

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Scalas, L. F. (2014). Doubly Latent Multilevel Analyses of Classroom Climate: An Illustration. *The Journal of Experimental Education*, *82*(2), 143–167. http://doi.org/10.1080/00220973.2013.769412

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA USA and Amsterdam, the Netherlands: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://timssandpirls.bc.edu/timss2011/reports/downloads/T11_IR_M_FrontMatter.pdf

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timss.bc.edu/PDF/t03_download/T03INTLMATRPT.pdf

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 International Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://pirls.bc.edu/PDF/PIRLS2006_international_report.pdf

Mullis, I. V. S., Martin, M. O., Robitaille, D. F., & Foy, P. (2009). *TIMSS Advanced 2008. International Report. Findings from IEA's Study of Achievement in Advanced Mathematics and Physics in the Final Year of Secondary School*. Chestnut Hill, MA: TIMSS & PRLS International Study Center, Boston College. Retrieved from http://pirls.bc.edu/timss_advanced/downloads/TA08_International_Report.pdf

Murnane, R. J., Singer, J. D., & Willett, J. B. (1988). Implications for Teacher Supply and Methodological Lessons for Research. *Educational Researcher*, *17*(6), 22–30. http://doi.org/10.3102/0013189X017006022

Murnane, R. J., & Willett, J. B. (2011). *Methods Matter: Improving causal Inference in Educational and Social Science Research*. Oxford: Oxford University Press.

Murray, H. A. (1938). *Explorations in personality; a clinical and experimental study of fifty men of college age*. New York: Oxford University Press.

Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S.-T., Yang, C.-C., … Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics*, *3*(4), 459–475. http://doi.org/10.1093/biostatistics/3.4.459

Muthen, B., & Masyn, K. (2005). Discrete-Time Survival Mixture Analysis. *Journal of Educational and Behavioral Statistics*, *30*(1), 27–58. http://doi.org/10.3102/10769986030001027

Muthén, L. K., & Muthén, B. (2012). *Mplus User's Guide* (Seventh Ed). Los Angeles, CA: Muthén & Muthén. Retrieved from http://www.statmodel.com/

Nagengast, B., & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, *104*(4), 1033–1053. http://doi.org/10.1037/a0027697

Nakamoto, J., & Schwartz, D. (2010). Is peer victimization associated with academic achievement? A meta-analytic review. *Social Development*, *19*(2), 221–242. http://doi.org/10.1111/j.1467-9507.2009.00539.x

Napier, J. L., & Jost, J. T. (2008). The "Antidemocratic Personality" Revisited: A Cross-National Investigation of Working-Class Authoritarianism. *Journal of Social Issues*, *64*(3), 595–617. http://doi.org/10.1111/j.1540-4560.2008.00579.x

Newmann, F. M. (1991). Classroom Thoughtfulness and Students' Higher Order Thinking: Common Indicators and Diverse Social Studies Courses. *Theory & Research in Social Education*, *19*(4), 410–433. http://doi.org/10.1080/00933104.1991.10505649

Norwalk, K. E., Hamm, J. V., Farmer, T. W., & Barnes, K. L. (2015). Improving the School Context of Early Adolescence Through Teacher Attunement to Victimization: Effects on School Belonging. *The Journal of Early Adolescence*, 1–21. http://doi.org/10.1177/0272431615590230

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling*, *14*(4), 535–569. http://doi.org/10.1080/10705510701575396

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, *41*(5), 673–690. http://doi.org/10.1007/s11135-006-9018-6

O'Connell, A. A., & McCoach, D. B. (2008). *Multilevel Modeling of Educational Data*. Charlotte, NC: IAP.

OECD. (2005). *Teachers Matter: Attracting, Developing and Retaining Effective Teachers* (Vol. 31). Paris: OECD Publishing. http://doi.org/10.1787/9789264018044-en

OECD. (2010a). *PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices (Volume IV)*. Paris: OECD Publishing. http://doi.org/10.1787/9789264091559-en

OECD. (2010b). *TALIS 2008 Technical Report*. Paris: OECD Publishing. http://doi.org/10.1787/9789264079861-en

OECD. (2012). *PISA 2012 Results: Excellence Through Equity: Giving Every Student the Chance to Succeed*. *PISA 2012 Results* (Vol. II). Paris: OECD Publishing. http://doi.org/10.1787/9789264201132-en

OECD. (2013a). *Pensions at a Glance 2013*. Paris: OECD Publishing. http://doi.org/10.1787/pension_glance-2013-en

OECD. (2013b). *TALIS 2013 Technical Report*. Paris: OECD Publishing. Retrieved from http://www.oecd.org/edu/school/TALIS-technical-report-2013.pdf

OECD. (2014a). *Society at a Glance 2014: OECD Social Indicators*. Paris: OECD Publishing. http://doi.org/10.1787/soc_glance-2014-en

OECD. (2014b). TALIS 2013 England Headteacher Questionnaires. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323129/TALIS_2013_England_HT_Questionnaire.pdf

OECD. (2014c). TALIS 2013 England Teacher Questionnaires. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323128/TALIS_2013_England_Teacher_Questionnaire.pdf

OECD. (2014d). TALIS 2013 Questionnaires. Retrieved from http://www.oecd.org/edu/school/Questionnaires TALIS 2013.pdf

OECD. (2014e). *TALIS 2013 Results An International Perspective on Teaching and Learning*. Paris: OECD Publishing. http://doi.org/10.1787/9789264196261-en

OECD. (2014f). Technical notes on indices and analysis used in TALIS 2013. In *TALIS 2013 Results: An International Perspective on Teaching and Learning* (pp. 213–255). Paris: OECD Publishing. http://doi.org/10.1787/9789264196261-12-en

Ortúzar, M. S., Flores, C., Milesi, C., & Cox, C. (2009). Aspectos de la formación inicial docente y su influencia en el rendimiento académico de los alumnos. In *Camino al Bicentenario, propuestas para Chile* (pp. 155–186). Santiago de Chile: Pontificia Universidad Católica de Chile.

Osterman, K. F. (2000). Students' Need for Belonging in the School Community. *Review of Educational Research*, *70*(3), 323–367. http://doi.org/10.3102/00346543070003323

Ostroff, C. (1992). The relationship between satisfaction, attitudes, and performance: An organizational level analysis. *Journal of Applied Psychology*, *77*(6), 963–974. http://doi.org/10.1037/0021-9010.77.6.963

Pearl, J. (2012). The Causal Foundations of Structural Equation Modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 68–91). New York: Guilford Press.

Perry, A. C. (1908). *The management of a city school*. New York: Macmillan Co.

Petras, H., Masyn, K. E., Buckley, J. a., Ialongo, N. S., & Kellam, S. (2011). Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual- and context-level influences. *Journal of Educational Psychology*, *103*(1), 223–237. http://doi.org/10.1037/a0021545

Phillips, M. (1997). What Makes Schools Effective? A Comparison of the Relationships of Communitarian Climate and Academic Climate to Mathematics Achievement and Attendance During Middle School. *American Educational Research Journal*, *34*(4), 633–662. http://doi.org/10.3102/00028312034004633

Podgursky, M., Monroe, R., & Watson, D. (2004). The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review*, *23*(5), 507–518. http://doi.org/10.1016/j.econedurev.2004.01.005

Podsakoff, N. P., LePine, J. A., & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology*, *92*(2), 438–454. http://doi.org/10.1037/0021-9010.92.2.438

Pokropek, A. (2015). Phantom Effects in Multilevel Compositional Analysis: Problems and Solutions. *Sociological Methods & Research*, *44*(4), 677–705. http://doi.org/10.1177/0049124114553801

Ponzo, M. (2013). Does bullying reduce educational achievement? An evaluation using matching estimators. *Journal of Policy Modeling*, *35*(6), 1057–1078. http://doi.org/10.1016/j.jpolmod.2013.06.002

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227. http://doi.org/10.1080/00273170701341316

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011a). Alternative Methods for Assessing Mediation in Multilevel Data: The Advantages of Multilevel SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 161–182. http://doi.org/10.1080/10705511.2011.557329

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011b). Mplus syntax files for single- and multilevel mediation models. Retrieved from http://www.statmodel.com/download/Preacher.pdf

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2015). Multilevel Structural Equation Models for Assessing Moderation Within and Across Levels of Analysis. *Psychological Methods*. http://doi.org/10.1037/met0000052

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209–33. http://doi.org/10.1037/a0020141

Prieto, L. L., Soria, M. S., Martínez, I. M., & Schaufeli, W. (2008). Extension of the Job Demands-Resources model in the prediction of burnout and engagement among teachers over time. *Psicothema*, *20*(3), 354–360.

Primo, D. M., Jacobsmeier, M. L., & Milyo, J. (2007). Estimating the Impact of State Policies and Institutions with Mixed-Level Data. *State Politics & Policy Quarterly*, *7*(4), 446–459. http://doi.org/10.1177/153244000700700405

Quartz, K., Thomas, A., Anderson, L., Masyn, K., Barraza, K., & Olsen, B. (2008). Careers in motion: A longitudinal retention study of role changing among early-career urban educators. *The Teachers College Record*, *110*(1), 218–250. Retrieved from http://www.tcrecord.org/content.asp?contentid=14598

Quaynor, L. J. (2012). Citizenship education in Post-conflict contexts: A review of the literature. *Education, Citizenship and Social Justice*, *7*(1), 33–57. http://doi.org/10.1177/1746197911432593

Quiñones, M., Van den Broeck, A., & De Witte, H. (2012). Burnout in Chilean public teachers: analysis of the health impairment process. *Revista Chilena de Salud Pública*, *16*(3), 212–224. Retrieved from http://www.revistasaludpublica.uchile.cl/index.php/RCSP/article/viewFile/23126/24474

Rabe-Hesketh, S., & Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *169*(4), 805–

827. http://doi.org/10.1111/j.1467-985X.2006.00426.x

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition* (3rd ed.). College Station, TX: Stata Press.

Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: MESA Press.

Reynolds, D. (1982). The Search for Effective Schools. *School Organisation*, *2*(3), 215–237. http://doi.org/10.1080/0260136820020302

Reynolds, D. (1997). School effectiveness: retrospect and prospect. *Scottish Educational Review*, *29*(2), 97–113. Retrieved from http://www.scotedreview.org.uk/media/scottish-educational-review/articles/51.pdf

Rivero, M. del R. (2013). *The Chilean Teacher Labor Market. University of California, Berkeley*. University of California, Berkeley. Retrieved from http://escholarship.org/uc/item/1jw6x1jn

Rivero, M. del R. (2015). The Link of Teacher Career Paths on the Distribution of High Qualified Teachers: A Chilean Case Study. *Archivos Analíticos de Políticas Educativas*, *23*(73). http://doi.org/10.14507/epaa.v23.1710

Rivero, M. del R., Hurtado, C., & San Martín, E. (2015). Exámenes de egreso , validez y concecuencias sobre los egresados : el caso chileno. In *COLMEE: Segundo Congreso Latinoamericano de Medición y Evaluación Educacional*. Mexico DF, Mexico. Retrieved from http://www.colmee.mx/public/conferences/1/presentaciones/ponenciasdia3/50Examenes.pdf

Roach, A. T., & Kratochwill, T. R. (2004). Evaluating School Climate and School Culture. *TEACHING Exceptional Children*, *37*(1), 10–17. http://doi.org/10.1177/004005990403700101

Robinson, J. P. (2014). Causal Inference and Comparative Analysis with Large-Scale Assessment Data. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of International large-scale assessment: background, technical issues, and methods of data analysis* (pp. 521–545). Boca Raton, London, New York: Chapman and Hall/CRC.

Rogers, W. T., Anderson, J. O., Klinger, D. A., & Dawber, T. (2006). Pitfalls and Potential of Secondary Data Analysis of the Council of Ministers of Education, Canada, National Assessment. *Canadian Journal of Education / Revue Canadienne de L'éducation*, *29*(3), 757. http://doi.org/10.2307/20054194

Román, M., & Murillo, J. (2011). Latin America: school bullying and academic achievement. *CEPAL Review*, (104), 37–53. Retrieved from http://www.cepal.org/publicaciones/xml/2/45332/RVI104RomanMurillo.pdf

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2012). How Teacher Turnover Harms Student Achievement. *American Educational Research Journal*, *50*(1), 4–36. http://doi.org/10.3102/0002831212463813

Rosenshine, B., & Furst, N. (1971). Research on teacher performance criteria. In B. O. Smith (Ed.), *Research in Teacher Education, a Symposium* (pp. 37–72). Englewood Cliffs, N.J.: PrenticeHall.

Rosenthal, R. (2002). The Pygmalion Effect and its Mediating Mechanisms. In J. M. Aronson (Ed.), *Improving Academic Achievement: Impact of Psychological Factors on Education*. Bingley, United States: Academic Press. http://doi.org/10.1016/B978-012064455-1/50005-1

Rubie-Davies, C. (2014). Teacher's Instructional Beliefs and The Classroom Climate: Connections and Conumdrums. In H. Fives & M. G. Gill (Eds.), *International Handbook of Research on Teachers' Beliefs* (pp. 266–283). New York and London: Routledge.

Ruffinelli, A., & Guerrero, A. (2009). Circulo de segmentacion del sistema educativo chileno: Destino laboral de los egresados de pedagogia en educacion basica. *Calidad En La Educacion*, (31), 20–44. Retrieved from http://www.cned.cl/public/secciones/SeccionRevistaCalidad/doc/64/cse_articulo832.pdf

Rutkowski, D., Rutkowski, L., & Wild, J. (2013). Predictors of School Violence Internationally: The Importance of Immigrant Status and Other Factors. In *5th IEA International Research Conference, 26–28 June 2013, Singapore*. Retrieved from http://www.iea.nl/fileadmin/user_upload/IRC/IRC_2013/Papers/IRC-2013_Rutkowski_etal.pdf

Rutkowski, L., Davier, M. von, & Rutkowski, D. (2013). *Handbook of International large-scale assessment: background, technical issues, and methods of data analysis*. (L. Rutkowski, M. von Davier, & D. Rutkowski, Eds.). Boca Raton, London, New York: Chapman and Hall/CRC.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher*, *39*(2), 142–151. http://doi.org/10.3102/0013189X10363170

Rutkowski, L., & Svetina, D. (2013). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. http://doi.org/10.1177/0013164413498257

Rutter, M., & Maughan, B. (2002). School Effectiveness Findings 1979–2002. *Journal of School Psychology*, *40*(6), 451–475. http://doi.org/10.1016/S0022-4405(02)00124-3

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1982). *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Harvard University Press.

Ryle, G. (1949). *The Concept of Mind* (60th Anniv). London and New York: Routledge.

Ryu, E. (2015). The Role of Centering for Interaction of Level 1 Variables in Multilevel Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(4), 617–630. http://doi.org/10.1080/10705511.2014.936491

Salamzadeh, Y. (2012). Role of emotionality and authoritarianism in unethical behaviors and their effect on organizational citizenship behavior: The case of fars gas company, Iran. *World Applied Sciences Journal*, *17*(4), 502–508. Retrieved from http://www.idosi.org/wasj/wasj17(4)12/15.pdf

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis. *Psychological Methods*, *8*(4), 448–467. http://doi.org/10.1037/1082-989X.8.4.448

Sandoval-Hernandez, A. (2008). School effectiveness research: a review of criticisms and some proposals to address them. *Educate~*, *1*(1), 31–44. Retrieved from http://educatejournal.org/index.php/educate/article/viewFile/141/152

Sandoval-Hernandez, A., Aghakasiri, P., Wild, J., & Rutkowski, D. (2013). Does increasing hours of schooling lead to improvements in student learning? *IEA's Policy Brief Series*, (1), 1–8. Retrieved from http://www.iea.nl/fileadmin/user_upload/Policy_Briefs/IEA_policy_brief_Sep2013.pdf

Santelices, M. V., & Taut, S. (2011). Convergent validity evidence regarding the validity of the Chilean standards-based teacher evaluation system. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 73–93. http://doi.org/10.1080/0969594X.2011.534948

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating Model Fit With Ordered Categorical Data Within a Measurement Invariance Framework: A Comparison of Estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(2), 167–180. http://doi.org/10.1080/10705511.2014.882658

Sauzet, O., Wright, K. C., Marston, L., Brocklehurst, P., & Peacock, J. L. (2013). Modelling the hierarchical structure in datasets with very small clusters: a simulation study to explore the effect of the proportion of clusters when the outcome is continuous. *Statistics in Medicine*, *32*(8), 1429–38. http://doi.org/10.1002/sim.5638

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. http://doi.org/10.1037/1082-989X.7.2.147

Scheerens, J. (1990). School Effectiveness Research and the Development of Process Indicators of School Functioning. *School Effectiveness and School Improvement*, *1*(1), 61–80. http://doi.org/10.1080/0924345900010106

Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement*, *24*(1), 1–38. http://doi.org/10.1080/09243453.2012.691100

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford, New York and Tokyo: Pergamon Press.

Scheerens, J., Glas, C., Thomas, S., & Thomas, S. (2003). *Educational evaluation, assessment and monitoring: A systematic approach*. Lisse, Abingdon, Exton (PA) and Tokyo: Swets & Zeitlinger Publishers.

Scheerens, J., Vermeulen, C., & Pelgrum, W. J. (1989). Generalizibility of instructional and school effectiveness indicators across nations. *International Journal of Educational Research*, *13*(7), 789–799. http://doi.org/10.1016/0883-0355(89)90029-3

Schein, E. H. (2004). *Organizational Culture and Leadership* (3rd ed.). San Francisco, CA: Jossey-Bass.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schneider, B., Ehrhart, M. G., & Macey, W. H. (2011). Organizational Climate Research. In N. M. Ashkanasy, C. P. M. Wilderom, & M. F. Peterson (Eds.), *The handbook of organizational culture and climate* (pp. 29–49). London EC1Y 1SP, UK: Sage Publications, Inc.

Schoen, L. T., & Teddlie, C. (2008). A new model of school culture: a response to a call for conceptual clarity. *School Effectiveness and School Improvement*, *19*(2), 129–153. http://doi.org/10.1080/09243450802095278

Schulz, W., Ainley, J., & Fraillon, J. (2011). *ICCS 2009 Technical Report*. (W. Schulz, J. Ainley, & J. Fraillon, Eds.). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Technical_Report.pdf

Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2010). *ICCS 2009 International Report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_International_Report.pdf

Schulz, W., Ainley, J., Friedman, T., & Lietz, P. (2011). *ICCS 2009 Latin American Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/fileadmin/user_upload/graphics/Publications/ICCS_2009_LA_Report.pdf

Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008). *International Civic and Citizenship Education Study: Assessment Framework.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Framework.pdf

Sibley, C. G., & Duckitt, J. (2008). Personality and Prejudice: A Meta-Analysis and Theoretical Review. *Personality and Social Psychology Review*, *12*(3), 248–279. http://doi.org/10.1177/1088868308319226

Singer, J. D., & Willett, J. B. (1988). Detecting Involuntary Layoffs in Teacher Survival Data: The Year of Leaving Dangerously. *Educational Evaluation and Policy Analysis*, *10*(3), 212–224. http://doi.org/10.3102/01623737010003212

Singer, J. D., & Willett, J. B. (1993). It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. *Journal of Educational and Behavioral Statistics*, *18*(2), 155–195. http://doi.org/10.3102/10769986018002155

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York: Oxford University Press.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, *75*(3), 417–453. http://doi.org/10.3102/00346543075003417

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic* *and advanced multilevel modeling* (2nd ed.). London: SAGE Publications Ltd.

Solvason, C. (2005). Investigating specialist school ethos … or do you mean culture? *Educational Studies*, *31*(1), 85–94. http://doi.org/10.1080/0305569042000310985

Son Hing, L. S., Bobocel, D. R., Zanna, M. P., & McBride, M. V. (2007). Authoritarian dynamics and unethical decision making: high social dominance orientation leaders and high right-wing authoritarianism followers. *Journal of Personality and Social Psychology*, *92*(1), 67–81. http://doi.org/10.1037/0022-3514.92.1.67

Sousa-Poza, A., & Henneberger, F. (2004). Analyzing Job Mobility with Job Turnover Intentions: An International Comparative Study. *Journal of Economic Issues*, *38*(1), 113–137. http://doi.org/10.1080/00213624.2004.11506667

Stancel-Piqtak, A., & Desa, D. (2014). Methodological Implementation of Multi Group Multilevel SEM with PIRLS2011: Improving Reading Achievement. In R. Strietholt, W. Bos, J.-E. Gustafsson, & M. Rosén (Eds.), *Educational Policy Evaluation Through International Comparative Assessments*. Münster: Waxmann Verlag GmbH.

Stapleton, L. M. (2002). The Incorporation of Sample Weights Into Multilevel Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 475–502. http://doi.org/10.1207/S15328007SEM0904_2

Stapleton, L. M. (2006). An Assessment of Practical Solutions for Structural Equation Modeling with Complex Sample Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(1), 28–58. http://doi.org/10.1207/s15328007sem1301_2

Stapleton, L. M. (2008). Variance Estimation Using Replication Methods in Structural Equation Modeling With Complex Sample Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(2), 183–210. http://doi.org/10.1080/10705510801922316

Stapleton, L. M. (2013). Incorporating Sampling Weights into Single- and Multilevel Analyses. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of International Largscale Assessment: background, technical issues, and methods of data analysis* (pp. 363–388). Boca Raton, London, New York: Chapman and Hall/CRC.

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Steel, R. P., & Lounsbury, J. W. (2009). Turnover process models: Review and synthesis of a conceptual literature. *Human Resource Management Review*, *19*(4), 271–282. http://doi.org/10.1016/j.hrmr.2009.04.002

Steel, R. P., & Ovalle, N. K. (1984). A review and meta-analysis of research on the relationship between behavioral intentions and employee turnover. *Journal of Applied Psychology*, *69*(4), 673–686. http://doi.org/10.1037//0021-9010.69.4.673

Sterba, S. K. (2015). A Latent Transition Analysis Model for Latent-State-Dependent Nonignorable Missingness. *Psychometrika*. http://doi.org/10.1007/s11336-015-9442-4

Stinebrickner, T. R. (1998). An empirical investigation of teacher attrition. *Economics of Education Review*, *17*(2), 127–136. http://doi.org/10.1016/S0272-

7757(97)00023-X

Stinebrickner, T. R. (2002). An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Leave. *The Journal of Human Resources*, *37*(1), 192–216. http://doi.org/10.2307/3069608

Stubager, R. (2008). Education effects on authoritarian–libertarian values: a question of socialization. *The British Journal of Sociology*, *59*(2), 327–350. http://doi.org/10.1111/j.1468-4446.2008.00196.x

Stubager, R. (2009). Education-based group identity and consciousness in the authoritarian-libertarian value conflict. *European Journal of Political Research*, *48*(2), 204–233. http://doi.org/10.1111/j.1475-6765.2008.00834.x

Swars, S. L., Meyers, B., Mays, L. C., & Lack, B. (2009). A Two-Dimensional Model of Teacher Retention and Mobility: Classroom Teachers and Their University Partners Take a Closer Look at a Vexing Problem. *Journal of Teacher Education*, *60*(2), 168–183. http://doi.org/10.1177/0022487108329116

Taki, M. (2009). Relations Among Bullying, Stresses, and Stressors. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of Bullying in Schools: An International Perspective* (pp. 151–162). Routledge.

Tan, X., Liu, L., Zheng, W., & Huang, Z. (2015). Effects of social dominance orientation and right-wing authoritarianism on corrupt intention: The role of moral outrage. *International Journal of Psychology*. http://doi.org/10.1002/ijop.12148

Taut, S., Santelices, M. V., & Stecher, B. (2012). Validation of a National Teacher Assessment and Improvement System. *Educational Assessment*, *17*(4), 163–199. http://doi.org/10.1080/10627197.2012.735913

Taut, S., Santelices, V., Araya, C., & Manzi, J. (2010). Theory underlying a national teacher evaluation program. *Evaluation and Program Planning*, *33*(4), 477–86. http://doi.org/10.1016/j.evalprogplan.2010.01.002

Televantou, I., Marsh, H. W., Kyriakides, L., Nagengast, B., Fletcher, J., & Malmberg, L.-E. (2015). Phantom effects in school composition research: consequences of failure to control biases due to measurement error in traditional multilevel models. *School Effectiveness and School Improvement*, *26*(1), 75–101. http://doi.org/10.1080/09243453.2013.871302

Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analyses based on meta-analytical findings. *Personnel Psychology*, *46*(2), 259–293. http://doi.org/10.1111/j.1744-6570.1993.tb00874.x

Thapa, A., Cohen, J., Guffey, S., & Higgins-D’Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research*, *83*(3), 357–385. http://doi.org/10.3102/0034654313483907

Thelen, H. A. (1950). Educational dynamics: Theory and research. *Journal of Social Issues*, *6*(2), 5–95.

Thomas, D. S. (1929). *Some new techniques for studying social behavior*. New York: Bureau of Publications, Teachers College, Columbia University.

Thrupp, M. (2001). Sociological and Political Concerns about School Effectiveness

Research: Time for a New Research Agenda. *School Effectiveness and School Improvement*, *12*(1), 7–40. http://doi.org/10.1076/sesi.12.1.7.3459

Torney-Purta, J., Amadeo, J., & Pilotti, F. (2004). *Fortalecimiento de la democracia en las Américas a través de la educación cívica: un análisis empírico que destaca las opiniones de los estudiantes y los maestros.* Washington D.C., United States: Organización de los Estados Americanos (OEA), Unidad de Desarrollo Social y Educación. Retrieved from http://www.educadem.oas.org/english/contenidos/strengthening democracy (IEA).pdf

Torres Irribarra, D., & Freund, R. (2014). Wright map: IRT item-person map with ConQuest integration. Retrieved from http://github.com/david-ti/wrightmap

Townsend, L., Flisher, A. J., Chikobvu, P., Lombard, C., & King, G. (2008). The Relationship between Bullying Behaviours and High School Dropout in Cape Town, South Africa. *South African Journal of Psychology, 38*(1), 21–32. http://doi.org/10.1177/008124630803800102

Ttofi, M. M., Farrington, D. P., & Lösel, F. (2012). School bullying as a predictor of violence later in life: A systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior*, *17*(5), 405–418. http://doi.org/10.1016/j.avb.2012.05.002

UNESCO. (2013). *Respuestas del Sector de Educación Frente al Bullying Homofóbico*. Santiago, Chile: OREALC/UNESCO Santiago. Retrieved from http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Havana/pdf/Educacion_bullying.pdf

Valenzuela, J. P. (2013). Profesores efectivos en todas las escuelas del país: Atraer, Formar y Retener. In *Profesión Docente en Chile: El (f)actor indispensable Profesores*. Pontificia Universidad Católica, Santiago, Chile, Santiago, Chile. Retrieved from http://www.ciae.uchile.cl/download.php?file=otros/2013CiclosDialogo/JP_Valenzuela.pdf

Valenzuela, J. P., Bellei, C., & Ríos, D. D. L. (2014). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy*, *29*(2), 217–241. http://doi.org/10.1080/02680939.2013.806995

Valenzuela, J. P., & Sevilla, A. (2013). *La movilidad de los nuevos profesores Chilenos en la década del 2000: un sistema escolar viviendo en peligro*. Santiago, Chile: Fondecyt Nº 1120740. Retrieved from http://www.ciae.uchile.cl/download.php?file=2015-docentes/Movilidad_Valenzuela y Sevilla.pdf

van Breukelen, W., van der Vlist, R., & Steensma, H. (2004). Voluntary employee turnover: combining variables from the“traditional” turnover literature with the theory of planned behavior. *Journal of Organizational Behavior, 25*(7), 893–914. http://doi.org/10.1002/job.281

van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134–150. http://doi.org/10.1016/j.edurev.2010.02.001

van Hiel, A., Pandelaere, M., & Duriez, B. (2004). The Impact of Need for Closure on

Conservative Beliefs and Racism: Differential Mediation by Authoritarian Submission and Authoritarian Dominance. *Personality and Social Psychology Bulletin*, *30*(7), 824–837. http://doi.org/10.1177/0146167204264333

van Horn, M. L. (2003). Assessing the Unit of Measurement for School Climate through Psychometric and Outcome Analyses of the School Climate Survey. *Educational and Psychological Measurement*, *63*(6), 1002–1019. http://doi.org/10.1177/0013164403251317

van Houtte, M. (2005). Climate or Culture? A Plea for Conceptual Clarity in School Effectiveness Research. *School Effectiveness and School Improvement*, *16*(1), 71–89. http://doi.org/10.1080/09243450500113977

Vivolo, A. M., Holt, M. K., & Massetti, G. M. (2011). Individual and Contextual Factors for Bullying and Peer Victimization: Implications for Prevention. *Journal of School Violence*, *10*(2), 201–212. http://doi.org/10.1080/15388220.2010.539169

Vollebergh. (1996). The development of authoritarianism in adolescence: Longitudinal change and the impact of age, gender and educational level. In K. Hurrelmann & S. Hamilton (Eds.), *Social problems and social contexts in adolescence: Perspectives across boundaries* (pp. 325–351). New York: Aldine de Gruyter.

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, *2*, 9–36. Retrieved from http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02_Chapter_01.pdf

Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology*, *59*(6), 414–419. http://doi.org/10.1037/h0026490

Wang, M.-T., & Degol, J. L. (2015). School Climate: a Review of the Construct, Measurement, and Impact on Student Outcomes. *Educational Psychology Review*. http://doi.org/10.1007/s10648-015-9319-1

West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear Mixed Models: A Practical Guide Using Statistical Software* (2nd ed.). Boca Raton: Chapman and Hall/CRC.

Westheimer, J., & Kahne, J. (2004). What Kind of Citizen? The Politics of Educating for Democracy. *American Educational Research Journal*, *41*(2), 237–269. http://doi.org/10.3102/00028312041002237

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, *91*(3), 461–481. http://doi.org/10.1037/0033-2909.91.3.461

Wiley, J. F., Bei, B., Trinder, J., & Manber, R. (2014). Variability as a Predictor: A Bayesian Variability Model for Small Samples and Few Repeated Measures. *arXiv Preprint*. Retrieved from http://arxiv.org/abs/1411.2961

Willett, J. B., & Singer, J. D. (1989). Two types of question about time: Methodological issues in the analysis of teacher career path data. *International Journal of Educational Research*, *13*(4), 421–437. http://doi.org/10.1016/0883-0355(89)90038-4

Willett, J. B., & Singer, J. D. (1991). From Whether to When: New Methods for

Studying Student Dropout and Teacher Attrition. *Review of Educational Research*, *61*(4), 407–450. http://doi.org/10.3102/00346543061004407

Willett, J. B., & Singer, J. D. (2004). Discrete Time Survival Analysis. In D. Kaplan (Ed.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (pp. 199–211). Thousand Oaks, California: Sage Publications, Inc.

Willms, J. D. (2010). School Composition and Contextual Effects on Student Outcomes. *Teachers College Record*, *112*(4), 1008–1037. Retrieved from https://www.tcrecord.org/Content.asp?ContentID=15658

Winter, N. (2008). SVR: Stata module to compute estimates with survey replication (SVR) based standard errors. *Statistical Software Components*. Retrieved from http://econpapers.repec.org/RePEc:boc:bocode:s427502

Withall, J. (1949). The Development of a Technique for the Measurement of Social-Emotional Climate in Classrooms. *The Journal of Experimental Education*, *17*(3), 347–361. http://doi.org/10.1080/00220973.1949.11010391

Wittgenstein, L. (2009). *Philosophical Investigations*. (P. Hacker & J. Schulte, Eds.) (4th ed.). Chichester, UK: Blackwell Publishing.

Woods, S., & Wolke, D. (2004). Direct and relational bullying among primary school children and academic achievement. *Journal of School Psychology*, *42*(2), 135–155. http://doi.org/10.1016/j.jsp.2003.12.002

Wormington, S. V., Anderson, K. G., Schneider, A., Tomlinson, K. L., & Brown, S. a. (2016). Peer Victimization and Adolescent Adjustment: Does School Belonging Matter? *Journal of School Violence*, *15*(1), 1–21. http://doi.org/10.1080/15388220.2014.922472

Wright, D. B. (1997). Extra-binomial variation in multilevel logistic models with sparse structures. *British Journal of Mathematical and Statistical Psychology*, *50*(1), 21–29. http://doi.org/10.1111/j.2044-8317.1997.tb01099.x

Wyatt, T. (1996). School effectiveness research: Dead end, damp squib or smouldering fuse? *Issues in Educational Research*, *6*(1), 79–112. Retrieved from http://www.iier.org.au/iier6/wyatt.html

Zamora, G. (2009). Compromisos organizacionales de los profesores chilenos y su relación con la intención de permanecer en sus escuelas. *Revista Latinoamericana de Psicologia*, *41*(3), 445–460. Retrieved from http://www.scielo.org.co/pdf/rlps/v41n3/v41n3a05

Zhao, N., Valcke, M., Desoete, A., & Verhaeghe, J. (2012). The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis: Implications for policies to foster education equity. *International Journal of Educational Development*, *32*(3), 412–422. http://doi.org/10.1016/j.ijedudev.2011.08.004

Zick, A., Küpper, B., Hövermann, A., German, C., & Fenn, E. (2011). *Intolerance , Prejudice and Discrimination A European Report*. Berin: Nora Langenbacher Friedrich-Ebert-Stiftung Forum Berlin Projekt. Retrieved from http://library.fes.de/pdf-files/do/07908-20110311.pdf

Zick, A., Wolf, C., Küpper, B., Davidov, E., Schmidt, P., & Heitmeyer, W. (2008). The Syndrome of Group-Focused Enmity: The Interrelation of Prejudices Tested with

Multiple Cross-Sectional and Panel Data. *Journal of Social Issues*, *64*(2), 363–383. http://doi.org/10.1111/j.1540-4560.2008.00566.x

Zullig, K. J., Koopman, T., & Huebner, E. (2009). Beyond GPA: Toward More Comprehensive Assessments of Students' School Experiences. *Child Indicators Research*, *2*(1), 95–108. http://doi.org/10.1007/s12187-008-9029-z

Zullig, K. J., Koopman, T. M., Patton, J. M., & Ubbes, V. A. (2010). School Climate: Historical Review, Instrument Development, and School Assessment. *Journal of Psychoeducational Assessment*, *28*(2), 139–152. http://doi.org/10.1177/0734282909344205

Zyphur, M. J., Zammuto, R. F., & Zhang, Z. (2016). Multilevel Latent Polynomial Regression for Modeling (In)Congruence Across Organizational Groups: The Case of Organizational Culture Research. *Organizational Research Methods*, *19*(1), 53–79. http://doi.org/10.1177/1094428115588570

Appendixes

Appendix (Paper 1)

Study 1: Selected Models compared to alternative model specifications

Discrete Time Survival Models is a flexible model which can be fitted in different ways (Muthen & Masyn, 2005; Willett & Singer, 2004). Within the latent variable framework, it is possible to easily relax the proportional hazard assumption, and allow covariates to have time variant effects. These models also are saturated models, and have no error term or variance. However, within this framework, a non-parametric frailty can be added to the model and we can thus account for non-observed heterogeneity of effects (Masyn, 2003, 2014).

As a means of **sensitivity analysis**, we estimated these alternative model specifications in comparison with our presented models. We rely on the Bayesian Information Criterion (BIC) which permits us to compare non nested models (Muthen & Masyn, 2005), is less computing intensive, and behaves well to identify mixture model miss-specifications under high sample size (n>1000) (Nylund, Asparouhov, & Muthén, 2007). The results of this comparison favours the models without a frailty term and with proportional odds. Therefore, we present the estimates of the most parsimonious model specification.

Table 1.11 Sensitivity Analysis

	Model 1		Model 2	
	BIC	df	BIC	df
Null Model	25054.91	6	21855.88	6
Selected Model	**24896.45**	22	**21704.17**	24
Frailty (non parametric)	24920.01	25	21730.49	27
Non Proportional odds	25424.22	102	22360.36	114

Notes: BIC = Bayesian Information Criterion, df = degrees of freedom or estimated parameters.

Study 2: Self report Measures

Study 2, uses self-report measures, answered by the teachers, in a complementary survey during the NTES process. In the following section, the original items in Spanish are displayed, along with a translation in English.

Table 1.12 Self report measures from Paper 2, Study 2

Principal leadership	
Cuando he enfrentado una dificultad en mi trabajo, he recibido el apoyo del Director para resolverla	When I've face a difficulty in my job, I've receive the support of the head teacher to solve it
El equipo directivo utiliza métodos efectivos para informarse sobre el desempeño de los profesores	The head management team uses effective methods to get information of from the performance of teachers
Los objetivos que busca alcanzar el establecimiento son conocidos por todos los profesores	The aims of the school are known by all teachers
Positive interpersonal relations	
Los profesores de esta escuela son respetados por la comunidad educativa	The teachers in the school are respected by the whole school community
Profesores y alumnos se llevan bien en esta escuela	Teachers and students get a long in this school
Academic Monitoring	
Desempeño en la evaluación docente	Performance in the teacher evaluation
Puntaje de los alumnos en prueba SIMCE	Students' scores in SIMCE test
Tasa de aprobación de los alumnos	Rate of students approval
School relations priority	
La relación del profesor con sus estudiantes	The relationship of the teacher with the students
La relación del docente con colegas y equipo directivo	The relationship with their colleagues and school head management staff

Notes: in left column Likert type items are displayed, and in the right column is a translation into English. Constructed scales titles are displayed in bold before each collection of items.

Appendix (Paper 2)

Items and scales were created using the original questions from the TALIS 2013 questionnaires for teachers and for school principals. Originally, these were applied in Spanish to the Chilean participants of this study. In this supplementary material, we employed the questions from the English version of the survey to facilitate the access to the readers and ease replication. In the present document, we copy portions of the questionnaires from Teachers (OECD, 2014c) and School Principals (i.e. Head teachers in England) (OECD, 2014b), thus displaying the original form of each question. Data and questionnaires can be access on the TALIS 2013 website (OECD, 2014d).

Measurement models were estimated for indicators of school intake, teachers' job satisfaction, and turnover intentions. We follow Desa (2014), and CFAs were conducted using the weights sample design, and school as cluster. Because stratification variables were not made available in the data release of TALIS 2013, we used a sandwich estimator to get correct standard errors. This a design based method, which accounts for clustering and protects for Type I error, by correcting standard error estimation with how much clustering is informative (McNeish & Stapleton, 2015; McNeish, 2014; Muthen & Masyn, 2005). Unavailability of stratification comes mainly at cost of higher standard errors, and it does not affects point estimates (Heeringa et al., 2009; Stapleton, 2006, 2008). Thus, using empirical clustered errors or sandwich estimator with survey weights is a more conservative test for measurement models at the population level, than Taylor Linearization and BRR method.

In general, we chose WLSMV or MLR estimator for categorical responses, depending on model fit, given their estimates vary in trustworthiness depending on the non-normality of the observed responses (Li, 2015; Sass, Schmitt, & Marsh, 2014).

In the next section, we review each variable of the current study and its preparation for analysis.

Measures and Items

Intake

Teachers answer the following questions:

Items 2A.1: School Intake

35. We would like to understand the composition of the target class. Please estimate the broad percentage of students who have the following characteristics.

'Socioeconomically disadvantaged homes' refers to homes with children eligible for Free School Meals.

This question asks about your personal perception of student background. It is acceptable to base your replies on rough estimates.

Students may fall into multiple categories.

Please mark one choice in each row.

	None	1% to 10%	11% to 30%	31% to 60%	More than 60%
a) Students whose first language is not English	☐1	☐2	☐3	☐4	☐5
b) Low academic achievers	☐1	☐2	☐3	☐4	☐5
c) Students with special needs	☐1	☐2	☐3	☐4	☐5
d) Students with behavioural problems	☐1	☐2	☐3	☐4	☐5
e) Students from socioeconomically disadvantaged homes (eligible for Free School Meals)	☐1	☐2	☐3	☐4	☐5
f) Academically gifted students	☐1	☐2	☐3	☐4	☐5

School intake was modelled as a single factor including items b, c, d, e. It was preferred to the inclusion of each of these characteristics as a single variable, given all of these were highly correlated (between .61 and .41 among all items). To create a single score, we fit a confirmatory factor analysis (CFA), modelling responses as ordinal answers, using the WLMSV estimator, while fixing the factor score variance to one, and freeing all factor loadings. We then retrieve the implied factor scores for each teacher. The measurement model yielded a good fit ($\chi^2(2)=5,00$, p=.08, CFI=.99, WRMR=.26). The WLMSV estimator was chosen given the estimates presented a close to normal distributions and this yields more trustworthy estimates (Sass et al., 2014) than other estimators under this scenario (Li, 2015).

School Administration

School Principals answer the following questions:

Items 2A.2: School Administration

10. Is this school publicly or privately managed?

Please mark one choice.

☐ 1 Publicly-managed
This is a school managed by a local education or government authority; e.g. community, foundation, VA and VC schools.

☐ 2 Privately-managed
This is a school managed by a non-government organisation; e.g. academies, free schools and independent schools).

11. Thinking about the funding of this school in a typical year, which of the following applies?

Please mark one choice in each row.

		Yes	No
a)	50% or more of the school's funding comes from central or local government.	☐ 1	☐ 2
b)	Teaching personnel are funded by central or local government.	☐ 1	☐ 2

These two questions were used to create the three main categories of school funding from Chile: public, subsidized, and private. Private schools are privately managed and do not receive any funding from the government. A school is public, if its publicly managed, more than 50% of the school funding comes from the local or central government, and additionally, all teaching personnel are funded by the local government. A school is subsidized, if it is privately managed, yet can receive either 50% or more funding from the local or central government, or if teacher personnel is funded by the government, or both. With these rules, three dummy variables were created to represent each type of school administration from the Chilean educational system.

School Location

School Principals answer the following questions:

Items 2A.3: School Location

9. Which best describes this school's location?

Please mark one choice.

- ☐1 Hamlet or rural area (1,000 people or fewer)
- ☐2 Village (1,001 to 3,000 people)
- ☐3 Small town (3,001 to 15,000 people)
- ☐4 Town (15,001 to 100,000 people)
- ☐5 City (100,001 to 1,000,000 people)
- ☐6 Large city (more than 1,000,000 people)

The answer to this question was ordered from one to six, from rural to urban. Thus, it comprises an ordinal measure of urbanity of the school location.

School Size

School Principals answer the following questions:

Items 2A.4: School Size

12. For each type of position listed below, please indicate the number of staff (head count) currently working in this school.

Staff may fall into multiple categories.
Please write a number in each row. Write 0 (zero) if there are none.

a) ⎿⏊⏊⏌ Teachers, irrespective of the grades/ages they teach
Those whose main professional activity at this school is the provision of instruction to students

b) ⎿⏊⏊⏌ Personnel for pedagogical support, irrespective of the grades/ages they support
Including all teacher aides or other non-teaching professionals who provide instruction or support teachers in providing instruction, professional curriculum/instructional specialists, educational media specialists, psychologists and nurses

c) ⎿⏊⏊⏌ School administrative personnel
Including receptionists, secretaries and administration assistants

d) ⎿⏊⏊⏌ School management personnel
Including headteachers, deputy/assistant headteachers, and other management staff whose main activity is management

e) ⎿⏊⏊⏌ Other staff

As a measure of school size, we used the first alternative of this question. This gives us a total head count of teachers within the teaching role for each school.

Age

Items 2A.5: Teachers Age

2. How old are you?

Please write a number.

|__|__| Years

Teachers indicated how old they were in years.

Experience

Items 2A.6: Teachers Experience

5. How many years' experience do you have?

Please round up to whole years.

a) |__|__| Year(s) working as a teacher at this school

b) |__|__| Year(s) working as a teacher in total

c) |__|__| Year(s) working in other education roles (do not include years working as a teacher)

d) |__|__| Year(s) working in other jobs

Teachers indicated how many years they have worked as a teacher (Item b). These answers were recorded as teacher labour experience.

Sex

Items 2A.7: Teachers Sex

1. Are you female or male?

☐ 1 Female

☐ 2 Male

Teachers indicated if they were either female or male. These answers were dummy coded, leaving males a reference category (female=1, male=0).

Teachers Educational qualifications

Items 2A.8: Teachers Certification

11. Have you completed a teacher training programme?

Please mark one choice. Select 'Yes' if you are currently on a Teach First programme.

- ☐1 Yes
- ☐2 No

Teachers indicated whether or not they had completed a teacher training program. Their answers were dummy coded, leaving teachers without teacher certification as a category of reference (yes=1, no=0).

Items 2A.9: Teachers Educational Qualifications

10. What is the highest level of formal education you have completed?

Please mark one choice.

- ☐1 GCE A levels or below, or equivalent
- ☐2 HNC, HND, NVQ at level 4+, Foundation Degree or equivalent
- ☐3 Bachelor's Degree
- ☐4 Master's Degree
- ☐5 Doctorate

Teachers indicated what was their last educational degree. These answers were dummy coded, and we left "Bachelor's Degree" as a University Degree for reference. The Chilean tertiary qualification for teachers last five years for a university degree, and is called licentiature, which is equivalent to at least four years of enrolment, and a professional degree comprising an additional 1 year of enrolment. Master degree, is considered a further qualification in Chile, which can only be undertaken after a licentiature degree.

Contract

Items2A.9: Type of Contract

6. What is your employment status as a teacher <u>at this school</u>?

Please mark one choice.

- ☐1 Permanent employment (an ongoing contract with no fixed end-point before the age of retirement)
- ☐2 Fixed-term contract for a period of more than 1 school year
- ☐3 Fixed-term contract for a period of 1 school year or less

Teachers indicated what kind of working contract they held with the current school. We dummy coded Full time contract, and left all other answers as category reference (full=1, else=0).

Works in more than one school

Items 2A.10: Works in more than one school

7. Do you currently work as a teacher of Key Stage 3 (age 11-14) at any other schools?

Please mark one choice.

☐[1] Yes

☐[2] No → **Please go to Question 9.**

Teachers answered if they worked in other schools besides the current school. These answers were dummy coded, leaving "no" as a reference category (yes=1, no=0).

Distributed Leadership

School Principals answer the following questions:

Items 2A.11: Distributed Leadership

22. How strongly do you agree or disagree with these statements as applied to this school?

Please mark one choice in each row.

	Strongly disagree	Disagree	Agree	Strongly agree
a) This school provides staff with opportunities to participate actively in school decisions.	☐[1]	☐[2]	☐[3]	☐[4]
b) This school provides parents or guardians with opportunities to participate actively in school decisions.	☐[1]	☐[2]	☐[3]	☐[4]
c) This school provides students with opportunities to participate actively in school decisions.	☐[1]	☐[2]	☐[3]	☐[4]
d) I make the important decisions on my own.	☐[1]	☐[2]	☐[3]	☐[4]
e) There is a collaborative school culture which is characterised by mutual support.	☐[1]	☐[2]	☐[3]	☐[4]

School's principal rate the following items to indicate the level of distributed leadership in the school. Factor scores were estimated using items A, B and C, employing the MLR estimator and declaring answers as continuous measures. The factors scores are the maximum a posteriori method (see OECD, 2013b, p. 156). Items D and E were not included in the final score of distributed leadership. In the present study, we use the estimated factor scores release by the OECD. These are scale to an international mean of 10, and to a standard deviation of 2 points.

Instructional Leadership

School Principals answer the following questions:

Items 2A.12: Instructional Leadership

21. Please indicate how frequently you engaged in the following in this school during the last 12 months.

Please mark one choice in each row.

		Never or rarely	Sometimes	Often	Very often
a)	I collaborated with teachers to solve classroom discipline problems.	☐1	☐2	☐3	☐4
b)	I observed instruction in the classroom.	☐1	☐2	☐3	☐4
c)	I took action to support co-operation among teachers to develop new teaching practices.	☐1	☐2	☐3	☐4
d)	I took action to ensure that teachers take responsibility for improving their teaching skills.	☐1	☐2	☐3	☐4
e)	I took action to ensure that teachers feel responsible for their students' learning outcomes.	☐1	☐2	☐3	☐4
f)	I provided parents or guardians with information on the school and student performance.	☐1	☐2	☐3	☐4
g)	I checked for mistakes and errors in school administrative procedures and reports.	☐1	☐2	☐3	☐4
h)	I resolved problems with the lesson timetable in this school.	☐1	☐2	☐3	☐4
i)	I collaborated with headteachers from other schools.	☐1	☐2	☐3	☐4

Factor scores were estimated using items C, D and E, employing the MLR estimator and declaring answers as continuous measures. The factors scores are the maximum a posteriori method (see OECD, 2013b, p. 156). In the present study, we used the estimated factor scores release by the OECD. These are scaled to an international mean of 10, with a standard deviation of 2.

Mutual Respect

School Principals answer the following questions:

Items 2A.13: Mutual Respect

30. How strongly do you agree or disagree with these statements as applied to this school?

Please mark one choice in each row.

		Strongly disagree	Disagree	Agree	Strongly agree
a)	The school staff share a common set of beliefs about schooling/learning.	☐1	☐2	☐3	☐4
b)	There is a high level of co-operation between the school and the local community.	☐1	☐2	☐3	☐4
c)	School staff have an open discussion about difficulties.	☐1	☐2	☐3	☐4
d)	There is mutual respect for colleagues' ideas.	☐1	☐2	☐3	☐4
e)	There is a culture of sharing success.	☐1	☐2	☐3	☐4
f)	The relationships between teachers and students are good.	☐1	☐2	☐3	☐4

Factor scores were estimated using items C, D, E and F, employing the MLR estimator and declaring answers as continuous measures. The factor scores are the maximum a posteriori method (see OECD, 2013b, p. 156). In the present study, we use the estimated factor scores released by the OECD. These are scaled to an international mean of 10, with a standard deviation of 2.

Shortage

School Principals answer the following questions:

Items 2A.14: Shortage

31. Is this school's capacity to provide quality instruction currently hindered by any of the following issues?

Please mark one choice in each row.

	Not at all	Very little	To some extent	A lot
a) Shortage of qualified and/or well performing teachers	☐1	☐2	☐3	☐4
b) Shortage of teachers with competence in teaching students with special needs	☐1	☐2	☐3	☐4
c) Shortage of vocational teachers	☐1	☐2	☐3	☐4
d) Shortage or inadequacy of instructional materials (e.g. textbooks)	☐1	☐2	☐3	☐4
e) Shortage or inadequacy of computers for instruction	☐1	☐2	☐3	☐4
f) Insufficient Internet access	☐1	☐2	☐3	☐4
g) Shortage or inadequacy of computer software for instruction	☐1	☐2	☐3	☐4
h) Shortage or inadequacy of library materials	☐1	☐2	☐3	☐4
i) Shortage of support personnel	☐1	☐2	☐3	☐4
j) Shortage of middle leaders	☐1	☐2	☐3	☐4

School principals indicated if schools were affected by shortages. We took item A as a proxy for previous teacher shortage.

School Delinquency

School Principals answer the following questions:

Items 2A.15: School Delinquency

32. In this school, how often do the following occur?

Please mark one choice in each row.

By students in this school:	Never	Rarely	Monthly	Weekly	Daily
a) Arriving late at school	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
b) Absenteeism (include only unjustified absences)	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
c) Cheating	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
d) Vandalism and theft	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
e) Intimidation or verbal abuse among students (or other forms of non-physical bullying)	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
f) Physical injury caused by violence among students	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
g) Intimidation or verbal abuse of teachers or staff	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$
h) Use/possession of drugs and/or alcohol in school	$\square_1$	$\square_2$	$\square_3$	$\square_4$	$\square_5$

Factor scores were estimated using items D, E, F and G, employing the MLR estimator and declaring answers as continuous measures. The factors scores are the maximum a posteriori method (see OECD, 2013b, p. 156). In the present study, we used the estimated factor scores release by the OECD. These were scaled to an international mean of 10, with a standard deviation of 2.

Positive Teacher-Student relations

Teachers answer the following questions:

Items 2A.16: Positive Teacher-Student relations

45. How strongly do you agree or disagree with the following statements about what happens in this school?

Please mark one choice in each row.

		Strongly disagree	Disagree	Agree	Strongly agree
a)	In this school, teachers and students usually get on well with each other.	☐1	☐2	☐3	☐4
b)	Most teachers in this school believe that the students' well-being is important.	☐1	☐2	☐3	☐4
c)	Most teachers in this school are interested in what students have to say.	☐1	☐2	☐3	☐4
d)	If a student from this school needs extra assistance, the school provides it.	☐1	☐2	☐3	☐4

Factor scores were estimated using all items, employing the MLR estimator and declaring answers as continuous measures. The factor scores are the maximum a posteriori method (see OECD, 2013b, p. 156). In the present study, we used the estimated factor scores release by the OECD. These were scaled to an international mean of 10, with a standard deviation of 2.

Classroom Discipline

Teachers answer the following questions:

Items 2A.18: Classroom Discipline

41. How strongly do you agree or disagree with the following statements about this target class?

Please mark one choice in each row.

		Strongly disagree	Disagree	Agree	Strongly agree
a)	When the lesson begins, I have to wait quite a long time for students to quieten down.	☐1	☐2	☐3	☐4
b)	Students in this class take care to create a pleasant learning atmosphere.	☐1	☐2	☐3	☐4
c)	I lose quite a lot of time because of students interrupting the lesson.	☐1	☐2	☐3	☐4
d)	There is a lot of disruptive noise in this classroom.	☐1	☐2	☐3	☐4

Factor scores were estimated using all items, employing the MLR estimator and declaring answers as continuous measures. The factor scores are the maximum a posteriori method (see OECD, 2013b, p. 156). In the present study, we used the estimated factor scores release by the OECD. These are scaled to an international mean of 10, with a standard deviation of 2.

Job Satisfaction, Withdrawal Cognitions and Intentions to Quit

Teachers answer the following questions:

Items 2A.19: Job attitudes scale

46. We would like to know how you generally feel about your job. How strongly do you agree or disagree with the following statements?

Please mark one choice in each row.

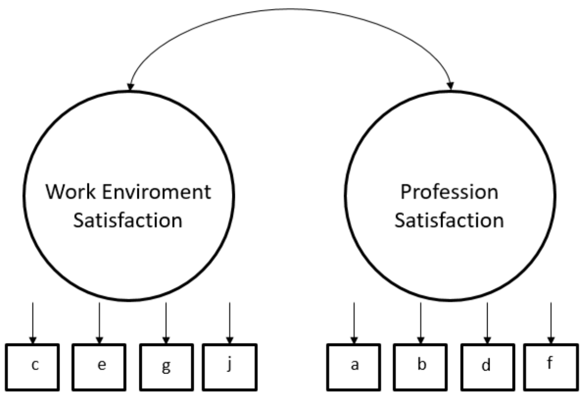
		Strongly disagree	Disagree	Agree	Strongly agree
a)	The advantages of being a teacher clearly outweigh the disadvantages.	□1	□2	□3	□4
b)	If I could decide again, I would still choose to work as a teacher.	□1	□2	□3	□4
c)	I would like to change to another school if that were possible.	□1	□2	□3	□4
d)	I regret that I decided to become a teacher.	□1	□2	□3	□4
e)	I enjoy working at this school.	□1	□2	□3	□4
f)	I wonder whether it would have been better to choose another profession.	□1	□2	□3	□4
g)	I would recommend my school as a good place to work.	□1	□2	□3	□4
h)	I think that the teaching profession is valued in society.	□1	□2	□3	□4
i)	I am satisfied with my performance in this school.	□1	□2	□3	□4
j)	All in all, I am satisfied with my job.	□1	□2	□3	□4

Originally these items were scaled together by the OECD (2013b, p. 206) as teacher job satisfaction, using the previously described method (MLR factor scores). This factor, comprise two additional sub factors within the scale: teachers' satisfaction with the current work environment (items C, E, G, J), and teachers' satisfaction with the profession (items A, B, D, F).

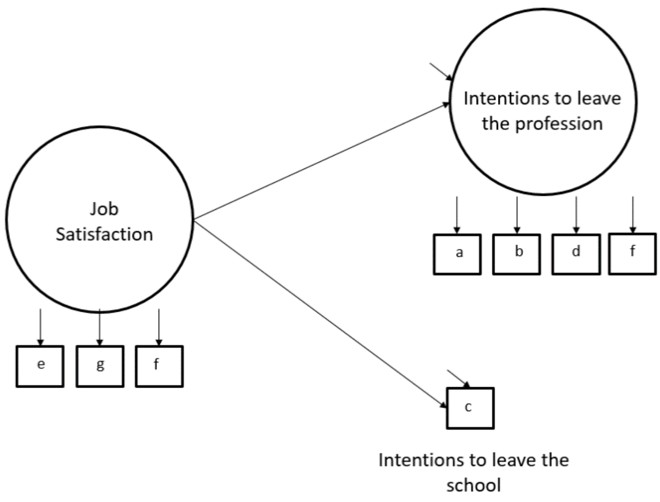
Because our interest is in withdrawal behaviours, we compared the measurement model described earlier (original model) with a measurement model in which the indicators were declared as categorical (categorical model); and finally we fitted a model in which withdrawal cognitions were separated as outcomes (alternative model). The original and the categorical model are essentially the same; what changes is the distributional assumption regarding the observed indicators. The original model assumes these are continuous, whereas the categorical model assumes answers to these indicators are categorical. The alternative also relies on the options of modelling indicators responses as categorical, and then goes one step further, and making explicit assumptions regarding the relationships between satisfaction and withdrawal cognitions. These models are depicted in Figure 2.2.

Figure 2.2 Diagram of endogenous variables

Original model



Alternative Model



In the alternative model, the item C, is conceptualized as a withdrawal cognition for intentions to leave the school. In contrast, items A, B, D, and F are withdrawal cognitions with respect to the profession (intentions to leave the profession). Finally, Items E, G and J are measures of Job Satisfaction, that is, how satisfied teachers feel working in the current school. In Table 6, we compare the model fit of the different modelling choices.

Table 2.7 Model Fit Comparison

Model	RMSEA	CFI	SRMR	χ2	df	SCF	p	BIC
Original	.10	.79	.08	295.60	19.00	1.82	.00	26538.04
Categorical	---	---	---	13613.63	65355.00	---	1.00	23910.32
Alternative	---	---	---	13616.24	65355.00	---	1.00	23910.32
Factors	---	---	---	7421.73	16285.00	---	1.00	20487.34

Note: RMSEA=root mean square error of approximation, CFI=comparative fit index, SRMR=standardized root mean square residuals, χ2=chi square, df=degrees of freedom, SCF=Scaling Correction Factor for MLR estimator, P=P value, BIC=Bayesian Information Criterion

The original model, in which two factors were fitted to explained the responses of items C, E, G, J, for Job Satisfaction and A, B, D and F for Professional Satisfaction, does not fit properly the observed responses in the Chilean population of teachers. The Categorical and Alternative model fit the data better. It reaches a non-significant χ^2, by modelling all response patterns. Essentially, the categorical model and the alternative model are not distinguishable by the model fit, given these have the same degrees of freedom. The main difference lies in the specification of item C, intentions to leave the school, predicted by the latent factor of Job Satisfaction. The BIC enables us to compare non-nested models, where the model with the lower BIC is deem preferable (Geiser, 2012; Heck & Thomas, 2015). Under this comparison, the alternative model is preferred over the original model (BIC alternative = 23910.321 < BIC original = 26538.036). The choice of the MLR estimator over the WLSMV estimator, for categorical indicators was preferred due to the observed non-normality of the items. For this scenario, MLR tends to outperform the WLSMV estimator (Li, 2015; Sass et al., 2014). Finally, “Factors” refers to the model used to compute the scores for each scale, in which only the indicators of teacher job satisfaction and teacher intentions to leave the profession were included.

To compute the factor scores for the revised version of Job Satisfaction and Intentions to leave the profession, we specified the items E, G, J, for Job Satisfaction and A, B, D and F for Withdrawal Cognition from the profession, and modelled their observed answers as categorical using the MLR estimator. The factor loadings were estimated freely, and the variances of each factor were fixed to one, thus producing factor scores standardized to the population. In the present study, we used these scores as measures, and intention to leave the school was preserved intact. Job satisfaction presents a mean of zero in the population CI95% [-.08, .08], where a higher score implies more satisfaction. Teacher intentions to leave the profession also present a mean of zero in the population with CI95%[-.07, .07], where higher scores indicate more withdrawal cognitions from the profession. Finally, item C was conserved intact from the release data, and scores are interpreted such that higher values are indicative of higher intentions to leave the present school.

Appendix (Paper 3)

MSEM estimates

Table 3.6 MSEM estimates, within covariates and random estimates

Parameter			Within E	Within (SE)		Between E	Between (SE)	
Intercepts		Math	--	--	--	412.00	(2.64)	*
		Engagement	--	--	--	-1.97	(2.72)	
		Bullying	--	--	--	.01	(.02)	
		Belonging	--	--	--	-.01	(.02)	
		Bullying * Belonging	--	--	--	-.05	(.01)	*
Sex (Female)	⇒	Math	-22.44	(2.85)	***	--	--	
	⇒	Engagement	-.25	(.07)	**	--	--	
	⇒	Belonging	.10	(.03)	***	--	--	
	⇒	Bullying	-.09	(.02)	***	--	--	
	⇒	Bullying * Belonging	.01	(.01)		--	--	
Younger	⇒	Math	-.63	(3.62)		--	--	
	⇒	Engagement	.12	(.09)		--	--	
	⇒	Belonging	.01	(.03)		--	--	
	⇒	Bullying	.07	(.03)	*	--	--	
	⇒	Bullying * Belonging	.01	(.02)		--	--	
Older	⇒	Math	-38.84	(3.27)	***	--	--	
	⇒	Engagement	-.14	(.07)		--	--	
	⇒	Belonging	-.09	(.03)	**	--	--	
	⇒	Bullying	.01	(.03)		--	--	
	⇒	Bullying * Belonging	.00	(.02)		--	--	
Residual Variances		Math	3204.24	(110.83)	***	478.82	(86.04)	*
		Engagement	2.37	(.09)	***	.06	(.22)	
		Bullying	.22	(.00)	***	.01	(.00)	*
		Belonging	.36	(.02)	***	.03	(.01)	*
		Bullying * Belonging	.10	(.01)	***	.00	(.00)	

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, ⇐ ⇒ expresses a covariance estimation.

Table 3.7 MSEM estimates, between covariates

Parameter			Within E	Within (SE)		Between E	Between (SE)	
Subsidized Schools	⇒	Math	--	--	--	4.30	(6.14)	
	⇒	Engagement	--	--	--	1.01	(1.52)	
	⇒	Belonging	--	--	--	.02	(.04)	
	⇒	Bullying	--	--	--	-.01	(.04)	
	⇒	Bullying * Belonging	--	--	--	.03	(.02)	
Private Schools	⇒	Math	--	--	--	33.50	(14.80)	*
	⇒	Engagement	--	--	--	.79	(1.71)	
	⇒	Belonging	--	--	--	.17	(.11)	
	⇒	Bullying	--	--	--	-.03	(.07)	
	⇒	Bullying * Belonging	--	--	--	.04	(.03)	
Safety	⇒	Math	--	--	--	1.99	(1.31)	
	⇒	Engagement	--	--	--	-.07	(.15)	
	⇒	Belonging	--	--	--	.03	(.01)	*
	⇒	Bullying	--	--	--	-.01	(.01)	
	⇒	Bullying * Belonging	--	--	--	.00	(.00)	
Discipline	⇒	Math	--	--	--	.20	(1.73)	
	⇒	Engagement	--	--	--	-.09	(.22)	
	⇒	Belonging	--	--	--	.01	(.01)	
	⇒	Bullying	--	--	--	-.02	(.01)	*
	⇒	Bullying * Belonging	--	--	--	.00	(.00)	
Academic Emphasis	⇒	Math	--	--	--	3.58	(1.33)	*
	⇒	Engagement	--	--	--	-.08	(.18)	
	⇒	Belonging	--	--	--	.02	(.01)	*
	⇒	Bullying	--	--	--	.00	(.01)	
	⇒	Bullying * Belonging	--	--	--	.00	(.00)	

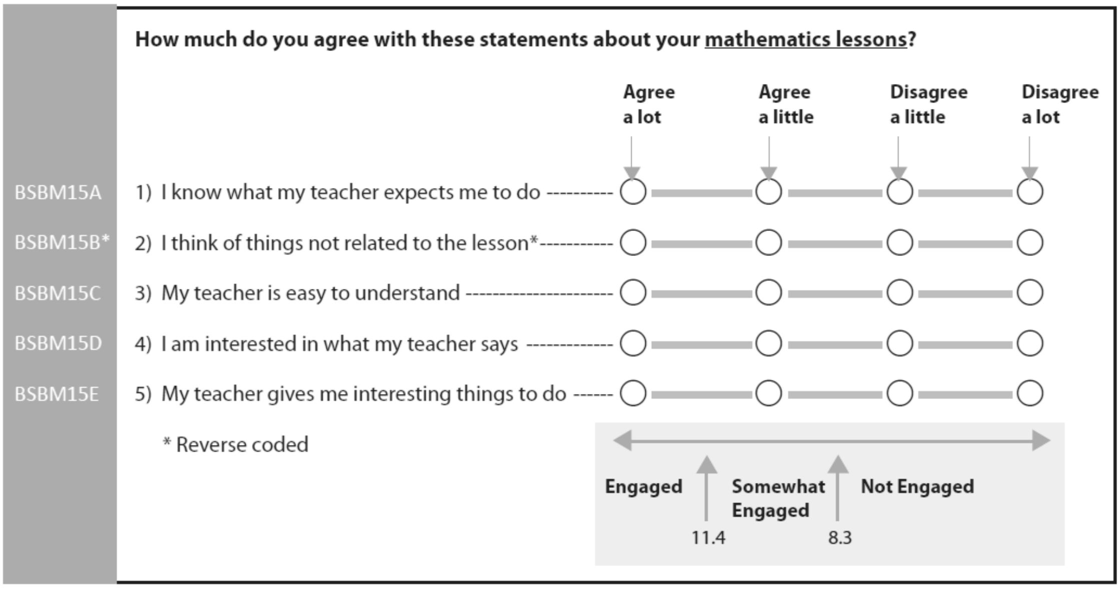
Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, ⇐ ⇒ expresses a covariance estimation.

Measures and Items

Engagement

Students indicated their level of agreement to the following items:

Items 3A.1: Engagement



How much do you agree with these statements about your mathematics lessons?

		Agree a lot	Agree a little	Disagree a little	Disagree a lot
BSBM15A	1) I know what my teacher expects me to do	○	○	○	○
BSBM15B*	2) I think of things not related to the lesson*	○	○	○	○
BSBM15C	3) My teacher is easy to understand	○	○	○	○
BSBM15D	4) I am interested in what my teacher says	○	○	○	○
BSBM15E	5) My teacher gives me interesting things to do	○	○	○	○

* Reverse coded

Students responses were modelled using a partial credit IRT, and the scale mean was set to 10, with a standard deviation of 2, for the international sample. Higher scores indicate a higher engagement.

Belonging

Students indicated their level of agreement to the following items:

Items 3A.2: Belonging

12

What do you think about your school? Tell how much you agree with these statements.

*Fill **one** circle for each line.*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I like being in school	○	○	○	○
b) I feel safe when I am at school	○	○	○	○
c) I feel like I belong at this school	○	○	○	○

Students answers were average, to create a composite score (alpha .69) to each student.

Bullying

Students indicated their level of agreement to the following items:

Items 3A.3: Bullying

During this year, how often have any of the following things happened to you at school?

		Never	A few times a year	Once or twice a month	At least once a week
BSBG13A	1) I was made fun of or called names	○	○	○	○
BSBG13B	2) I was left out of games or activities by other students	○	○	○	○
BSBG13C	3) Someone spread lies about me	○	○	○	○
BSBG13D	4) Something was stolen from me	○	○	○	○
BSBG13E	5) I was hit or hurt by other student(s) (e.g., shoving, hitting, kicking)	○	○	○	○
BSBG13F	6) I was made to do things I didn't want to do by other students	○	○	○	○

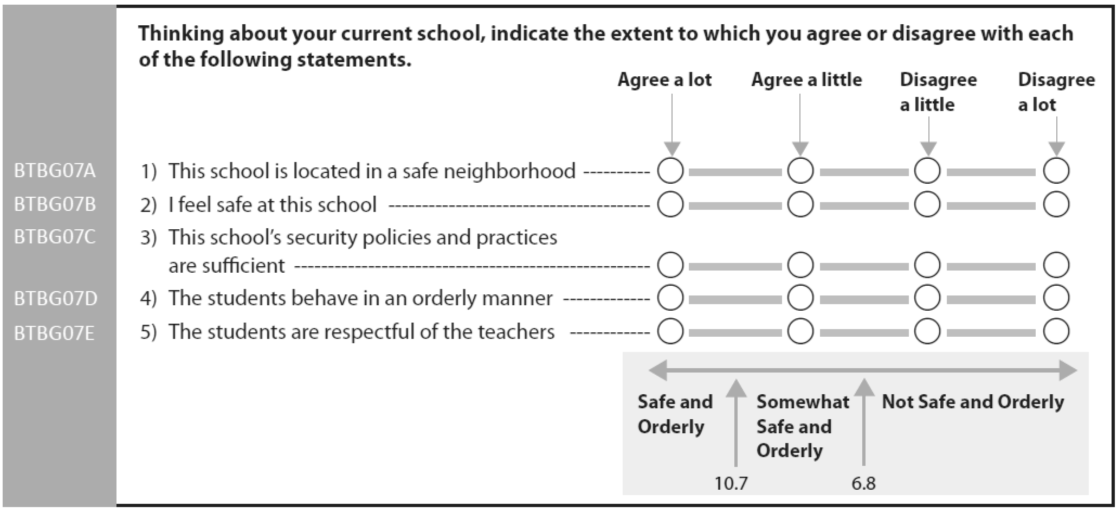
Almost Never | 9.6 | About Monthly | 7.7 | About Weekly

Students responses were modelled using a partial credit IRT, and the scale mean was set to 10, with a standard deviation of 2, for the international sample. In the current study. This scale was further discretized into three overall ordinal categories, as indicated in the previous figure. This, distinguishes those students who suffer some form of bullying with a weekly frequency, monthly, and almost never. In the present study we use this later form, to create a dummy variable. With this transformation we flagged all those students who suffered some form of bullying at least monthly, and left the rest of the students as a reference category.

School Safety

Math teachers from the target class group indicated the level of agreement to the following items:

Items 3A.4: School Safety



Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements.

		Agree a lot	Agree a little	Disagree a little	Disagree a lot
BTBG07A	1) This school is located in a safe neighborhood	○	○	○	○
BTBG07B	2) I feel safe at this school	○	○	○	○
BTBG07C	3) This school's security policies and practices are sufficient	○	○	○	○
BTBG07D	4) The students behave in an orderly manner	○	○	○	○
BTBG07E	5) The students are respectful of the teachers	○	○	○	○

Safe and Orderly | 10.7 | Somewhat Safe and Orderly | 6.8 | Not Safe and Orderly

Their answers where modelled with a partial credit IRT specification. With this measurement model, a score was created with a mean scale of 10, for the international sample, and standard deviation of 2. This is fixed covariate for schools. Higher score indicates safer schools.

School Discipline

School principal's indicated their level of agreement to the following items:

Items 3A.5: School Discipline

To what degree is each of the following a problem among eighth grade students in your school?

		Not a problem	Minor problem	Moderate problem	Serious problem
BCBG12AA	1) Arriving late at school	○	○	○	○
BCBG12AB	2) Absenteeism (i.e., unjustified absences)	○	○	○	○
BCBG12AC	3) Classroom disturbance	○	○	○	○
BCBG12AD	4) Cheating	○	○	○	○
BCBG12AE	5) Profanity	○	○	○	○
BCBG12AF	6) Vandalism	○	○	○	○
BCBG12AG	7) Theft	○	○	○	○
BCBG12AH	8) Intimidation or verbal abuse among students (including texting, emailing, etc.)	○	○	○	○
BCBG12AI	9) Physical injury to other students	○	○	○	○
BCBG12AJ	10) Intimidation or verbal abuse of teachers or staff (including texting, emailing, etc.)	○	○	○	○
BCBG12AK	11) Physical injury to teachers or staff	○	○	○	○

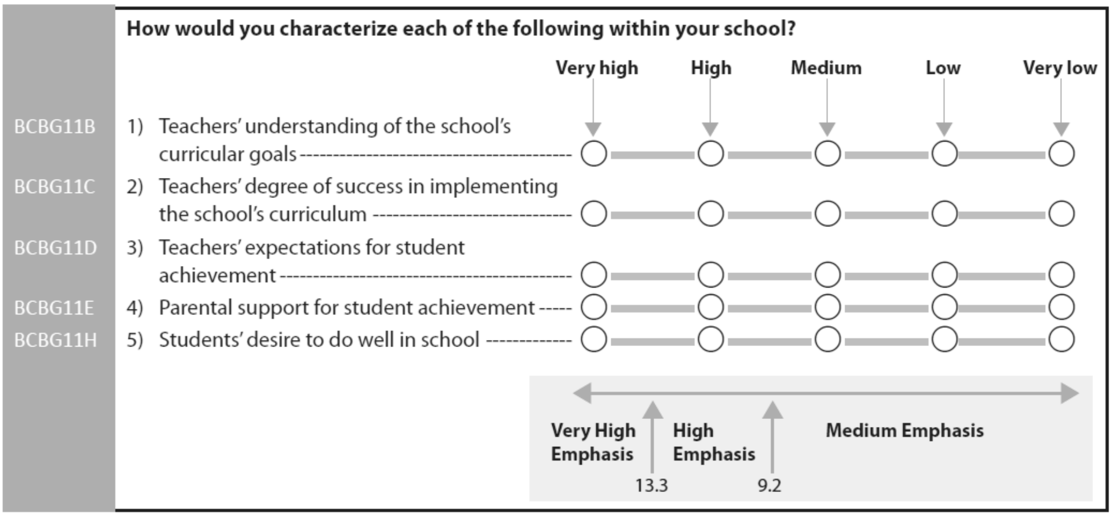
Hardly Any Problems | 10.7 | Minor Problems | 8.0 | Moderate Problems

Theirs answers were combined via a partial credit IRT model, and factors scores were derived. Similarly, to the previous scales, this score was scaled to an international mean of 10, and a standard deviation of 2. This is fixed covariate for schools. Higher scores indicate schools with more discipline.

School Academic emphasis

School principal's indicated their level of agreement to the following items:

Items 3A.6: Academic emphasis



How would you characterize each of the following within your school?

		Very high	High	Medium	Low	Very low
BCBG11B	1) Teachers' understanding of the school's curricular goals	○	○	○	○	○
BCBG11C	2) Teachers' degree of success in implementing the school's curriculum	○	○	○	○	○
BCBG11D	3) Teachers' expectations for student achievement	○	○	○	○	○
BCBG11E	4) Parental support for student achievement	○	○	○	○	○
BCBG11H	5) Students' desire to do well in school	○	○	○	○	○

Theirs answers were combined via a partial credit IRT model, in the same ways as previous scales (scores were scale at an international mean of 10, and a standard deviation of 2). Higher scores in this scale, indicate schools with a higher academic emphasis.

Appendix (Paper 4)

A) General Model: Equations

$yij = i1 + b1*mij + c1*xij + c2*wi + c3*wj + c4*xijwi + c5* xijwj + r1$

$mij = i2 + a1* xij + a2* wi + a3* wj + a4* xijwi + a5* xijwj + r2$

$xij = i3 + d1*zi + e1*zj + r3$

$wi = i4 + d2*zi + e2*zj + r4$

$wj = i5 + d3*zi + e3* zj + r5$

$xijwj = i6 + d4* zi + e4*zj + r6$

$xijwj = i7 + d5*z\_i + e5*zj + r7$

$cov2 = cov(xij,wi)$

$cov3 = cov(xij,wj)$

where,

i_1-i_7 = stand for intercepts of each equation

r_1-r_7 = stand for residuals of each equation

$_i$ = stand for within cluster deviations

$_j$ = stand for between cluster deviations

B) Total Effect Decomposition of "y on x"

Total effect decomposition of "y on x" accounting for between and within moderators' effects.

Indirect effect = (a1+a4*wi+a5*wj)*b1;

Pure Direct effect = c1 + c4*w_i + c5*w_j;

Total Effect = (a1+a4*wi+a5*wj)*b1 + c1 + c4 + c5;

Where,

wj = 0, due to prior centering of covariates at the grand mean

wi = 0, due to prior centering of covariates at the cluster mean

C) Conditional expected values of indirect effects

Expected values of indirect effects, at high and low values of the moderator, between effects.

Low = (a1+a5*lo)*b1;

Average = (a1+a5*me)*b1;

High = (a1+a5*hi)*b1;

Where,

lo = -10 (-1 SD), me = 0; hi = +10 (+1SD)

D) Linear Moderated Mediation Index formula (between effect)

LMM = a5*b1

E) Tables of parameters results for all the estimated Models

In the following tables (Tables 4A.2-4A.12) all estimated models are presented. These contain the unstandardized estimates for the presented model in Figure 4.2. Given the complexity of the models, we avoid the use of the name of the covariates. Instead, we use the referred term of the equations presented in Appendix A. In these tables, path coefficients are expressed by stating "$x_{ij} \Rightarrow y_{ij}$", which stands for the general linear model of the following form:

$$y_{ij} = \text{intercept} + x_{ij} + \ldots + \text{residual}$$

Each of these parameters has an index in the diagram presented in Figure 4.2, the following table expresses its equivalence.

Table 4A.1 Path coefficients equivalence with equation terms

Direct effects				Within Effects of SES				Intercepts	
c1	x_{ij}	⇒	y_{ij}	e1	z_j	⇒	x_{ij}	i1	y_{ij}
c2	w_i	⇒	y_{ij}	e2	z_j	⇒	w_i	i2	m_{ij}
c3	w_j	⇒	y_{ij}	e3	z_j	⇒	w_j	i3	x_{ij}
c4	$x_{ij}w_i$	⇒	y_{ij}	e4	z_j	⇒	$x_{ij}w_i$	i4	w_i
c5	$x_{ij}w_j$	⇒	y_{ij}	e5	z_j	⇒	$x_{ij}w_j$	i5	w_j
								i6	$x_{ij}w_i$
								i7	$x_{ij}w_j$

Effects on mediator				Between Effects of SES				Residuals	
a1	x_{ij}	⇒	m_{ij}	d1	z_i	⇒	x_{ij}	r1	y_{ij}
a2	w_i	⇒	m_{ij}	d2	z_i	⇒	w_i	r2	m_{ij}
a3	w_j	⇒	m_{ij}	d3	z_i	⇒	w_j	r3	x_{ij}
a4	$x_{ij}w_i$	⇒	m_{ij}	d4	z_i	⇒	$x_{ij}w_i$	r4	w_i
a5	$x_{ij}w_j$	⇒	m_{ij}	d5	z_i	⇒	$x_{ij}w_j$	r5	w_j
								r6	$x_{ij}w_i$
								r7	$x_{ij}w_j$

Mediator Effect				Covariance	
b1	m_{ij}	⇒	y_{ij}	cov2	$(x_{ij}, x_{ij}w_i)$
				cov3	$(x_{ij}, x_{ij}w_j)$

Notes: xij = civic Knowledge, mij= authoritarianism, yij= outcomes (gender equality, ethnic rights, permissibility of corruption, freedom of speech), z_{ij}=SES, w_{ij}= Open classroom discussion.

Table 4A.2 Path Analysis Estimates, Hong Kong SAR

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.23	(.03)	**	.21	(.03)	**	-.21	(.02)	**	.21	(.02)	**
c2	w_i	⇒	y_{ij}	.09	(.02)	**	.15	(.03)	**	-.04	(.02)	**	.07	(.02)	**
c3	w_j	⇒	y_{ij}	.16	(.08)		.24	(.09)	**	-.13	(.06)	*	.02	(.07)	
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		-.01	(.00)	*
c5	$x_{ij}w_j$	⇒	y_{ij}	.01	(.01)		.01	(.01)		.01	(.01)		.01	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.17	(.02)	**	-.08	(.03)	**	.38	(.02)	**	-.13	(.03)	**
a1	x_{ij}	⇒	m_{ij}	-.25	(.03)	**	-.25	(.03)	**	-.25	(.03)	**	-.25	(.03)	**
a2	w_i	⇒	m_{ij}	-.04	(.02)		-.03	(.02)		-.04	(.02)		-.03	(.02)	
a3	w_j	⇒	m_{ij}	.05	(.09)		.05	(.09)		.05	(.09)		.05	(.09)	
a4	$x_{ij}w_i$	⇒	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
a5	$x_{ij}w_j$	⇒	m_{ij}	-.01	(.01)		-.01	(.01)		-.01	(.01)		-.01	(.01)	
d1	z_i	⇒	x_{ij}	-.65	(.17)	**	-.65	(.17)	**	-.65	(.17)	**	-.65	(.17)	**
d2	z_i	⇒	w_i	-.02	(.27)		-.02	(.27)		-.02	(.27)		-.02	(.27)	
d3	z_i	⇒	w_j	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
d4	z_i	⇒	$x_{ij}w_i$	4.91	(4.86)		4.91	(4.86)		4.91	(4.86)		4.92	(4.86)	
d5	z_i	⇒	$x_{ij}w_j$	1.08	(.77)		1.08	(.77)		1.08	(.77)		1.08	(.77)	
e1	z_j	⇒	x_{ij}	5.27	(1.03)	**	5.27	(1.03)	**	5.27	(1.03)	**	5.27	(1.03)	**
e2	z_j	⇒	w_i	-.10	(.07)		-.11	(.07)		-.10	(.07)		-.10	(.07)	
e3	z_j	⇒	w_j	1.41	(.81)		1.41	(.81)		1.41	(.81)		1.41	(.81)	
e4	z_j	⇒	$x_{ij}w_i$	-2.59	(2.90)		-2.58	(2.90)		-2.58	(2.90)		-2.58	(2.90)	
e5	z_j	⇒	$x_{ij}w_j$	-5.93	(5.70)		-5.93	(5.70)		-5.93	(5.70)		-5.93	(5.70)	
	i1		y_{ij}	51.40	(.27)	**	52.01	(.25)	**	45.91	(.24)	**	47.15	(.29)	**
	i2		m_{ij}	.12	(.25)		.13	(.25)		.13	(.25)		.13	(.25)	
	i3		x_{ij}	-.04	(.60)		-.04	(.60)		-.04	(.60)		-.04	(.60)	
	i4		w_i	.07	(.04)		.08	(.04)		.07	(.04)		.07	(.04)	
	i5		w_j	.02	(.44)		.02	(.44)		.02	(.44)		.02	(.44)	
	i6		$x_{ij}w_i$	9.89	(1.98)	**	9.88	(1.99)	**	9.88	(1.99)	**	9.88	(1.98)	**
	i7		$x_{ij}w_j$	14.99	(2.73)	**	14.99	(2.73)	**	14.99	(2.73)	**	14.99	(2.73)	**
	r1		y_{ij}	78.18	(2.03)	**	102.44	(2.80)	**	65.00	(2.53)	**	89.52	(2.32)	**
	r2		m_{ij}	71.37	(2.88)	**	71.36	(2.87)	**	71.34	(2.88)	**	71.36	(2.87)	**
	r3		x_{ij}	85.40	(5.16)	**	85.40	(5.16)	**	85.40	(5.16)	**	85.40	(5.16)	**
	r4		w_i	96.00	(4.97)	**	95.99	(4.97)	**	96.00	(4.97)	**	95.99	(4.97)	**
	r5		w_j	14.07	(1.80)	**	14.07	(1.80)	**	14.07	(1.80)	**	14.07	(1.80)	**
	r6		$x_{ij}w_i$	11296.33	(1124.04)	**	11296.37	(1124.00)	**	11296.33	(1123.99)	**	11296.35	(1124.00)	**
	r7		$x_{ij}w_j$	1429.42	(199.51)	**	1429.42	(199.51)	**	1429.42	(199.51)	**	1429.42	(199.51)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	10.10	(1.90)	**	10.10	(1.90)	**	10.09	(1.91)	**	10.09	(1.90)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	13.12	(2.37)	**	13.12	(2.37)	**	13.12	(2.37)	**	13.12	(2.37)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.3 Path Analysis Estimates, Indonesia

Label	Path			Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
				E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.27	(.02)	**	.29	(.03)	**	-.33	(.02)	**	.39	(.03)	**
c2	w_i	⇒	y_{ij}	.07	(.01)	**	.12	(.01)	**	-.06	(.01)	**	.12	(.02)	**
c3	w_j	⇒	y_{ij}	.10	(.03)	**	.20	(.05)	**	-.19	(.05)	**	.28	(.05)	**
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)	*	.00	(.00)		.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	⇒	y_{ij}	.01	(.00)		.01	(.01)	*	.01	(.01)		.00	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.06	(.02)	**	.20	(.03)	**	.41	(.02)	**	.17	(.03)	**
a1	x_{ij}	⇒	m_{ij}	-.31	(.03)	**	-.31	(.03)	**	-.31	(.03)	**	-.31	(.03)	**
a2	w_i	⇒	m_{ij}	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
a3	w_j	⇒	m_{ij}	-.08	(.04)		-.08	(.04)		-.08	(.04)		-.08	(.04)	
a4	$x_{ij}w_i$	⇒	m_{ij}	-.01	(.00)	*	-.01	(.00)	*	.00	(.00)	*	-.01	(.00)	*
a5	$x_{ij}w_j$	⇒	m_{ij}	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**
d1	z_i	⇒	x_{ij}	.30	(.11)	**	.30	(.11)	**	.30	(.11)	**	.30	(.11)	**
d2	z_i	⇒	w_i	.29	(.19)		.29	(.19)		.29	(.19)		.29	(.19)	
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	1.15	(1.18)		1.15	(1.18)		1.15	(1.18)		1.16	(1.18)	
d5	z_i	⇒	$x_{ij}w_j$	.07	(.47)		.07	(.47)		.07	(.47)		.07	(.47)	
e1	z_j	⇒	x_{ij}	4.14	(.67)	**	4.14	(.67)	**	4.14	(.67)	**	4.14	(.67)	**
e2	z_j	⇒	w_i	-.02	(.01)		-.02	(.01)		-.02	(.01)		-.01	(.01)	
e3	z_j	⇒	w_j	2.47	(.38)	**	2.47	(.38)	**	2.47	(.38)	**	2.47	(.38)	**
e4	z_j	⇒	$x_{ij}w_i$	-1.78	(1.63)		-1.78	(1.63)		-1.78	(1.63)		-1.78	(1.63)	
e5	z_j	⇒	$x_{ij}w_j$	9.80	(2.64)	**	9.80	(2.64)	**	9.80	(2.64)	**	9.80	(2.64)	**
	i1		y_{ij}	42.21	(.15)	**	50.12	(.17)	**	53.22	(.18)	**	47.99	(.18)	**
	i2		m_{ij}	.19	(.17)		.20	(.17)		.19	(.17)		.19	(.17)	
	i3		x_{ij}	.04	(.32)		.04	(.32)		.04	(.32)		.04	(.32)	
	i4		w_i	-.01	(.01)		-.01	(.01)		-.01	(.01)		-.01	(.01)	
	i5		w_j	-.02	(.29)		-.02	(.29)		-.02	(.29)		-.02	(.29)	
	i6		$x_{ij}w_i$	9.30	(.89)	**	9.31	(.89)	**	9.31	(.89)	**	9.31	(.89)	**
	i7		$x_{ij}w_j$	8.26	(1.50)	**	8.26	(1.50)	**	8.26	(1.50)	**	8.26	(1.50)	**
	r1		y_{ij}	31.30	(1.31)	**	49.87	(1.34)	**	47.07	(1.29)	**	73.85	(2.00)	**
	r2		m_{ij}	34.36	(1.43)	**	34.36	(1.43)	**	34.36	(1.43)	**	34.35	(1.43)	**
	r3		x_{ij}	42.49	(2.78)	**	42.49	(2.78)	**	42.49	(2.78)	**	42.49	(2.78)	**
	r4		w_i	78.33	(2.59)	**	78.33	(2.58)	**	78.33	(2.59)	**	78.34	(2.59)	**
	r5		w_j	12.14	(1.74)	**	12.14	(1.74)	**	12.14	(1.74)	**	12.14	(1.74)	**
	r6		$x_{ij}w_i$	3686.78	(290.23)	**	3686.81	(290.24)	**	3686.80	(290.24)	**	3686.79	(290.24)	**
	r7		$x_{ij}w_j$	660.09	(101.34)	**	660.09	(101.34)	**	660.09	(101.34)	**	660.08	(101.34)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	9.31	(.89)	**	9.31	(.89)	**	9.31	(.89)	**	9.32	(.89)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	4.94	(1.31)	**	4.94	(1.31)	**	4.94	(1.31)	**	4.94	(1.31)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.4 Path Analysis Estimates, Republic of Korea

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	$\Rightarrow$	y_{ij}	.29	(.02)	**	.29	(.02)	**	-.19	(.02)	**	.52	(.02)	**
c2	w_i	$\Rightarrow$	y_{ij}	.03	(.01)	**	.05	(.02)	**	.00	(.01)		.04	(.02)	*
c3	w_j	$\Rightarrow$	y_{ij}	.19	(.06)	**	.27	(.07)	**	-.05	(.04)		.05	(.06)	
c4	$x_{ij}w_i$	$\Rightarrow$	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	$\Rightarrow$	y_{ij}	.00	(.01)		-.01	(.01)		.00	(.01)		-.01	(.01)	
b1	m_{ij}	$\Rightarrow$	y_{ij}	-.18	(.02)	**	-.15	(.02)	**	.40	(.01)	**	-.19	(.02)	**
a1	x_{ij}	$\Rightarrow$	m_{ij}	-.39	(.02)	**	-.39	(.02)	**	-.39	(.02)	**	-.39	(.02)	**
a2	w_i	$\Rightarrow$	m_{ij}	.01	(.02)		.01	(.02)		.01	(.02)		.01	(.02)	
a3	w_j	$\Rightarrow$	m_{ij}	.04	(.07)		.04	(.07)		.04	(.07)		.04	(.07)	
a4	$x_{ij}w_i$	$\Rightarrow$	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
a5	$x_{ij}w_j$	$\Rightarrow$	m_{ij}	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
d1	z_i	$\Rightarrow$	x_{ij}	2.29	(.16)	**	2.29	(.16)	**	2.29	(.16)	**	2.29	(.16)	**
d2	z_i	$\Rightarrow$	w_i	.48	(.15)	**	.48	(.15)	**	.48	(.15)	**	.48	(.15)	**
d3	z_i	$\Rightarrow$	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	$\Rightarrow$	$x_{ij}w_i$	-.70	(1.64)		-.70	(1.64)		-.70	(1.64)		-.71	(1.64)	
d5	z_i	$\Rightarrow$	$x_{ij}w_j$	.79	(.64)		.79	(.64)		.79	(.64)		.79	(.64)	
e1	z_j	$\Rightarrow$	x_{ij}	3.57	(.39)	**	3.57	(.39)	**	3.57	(.39)	**	3.57	(.39)	**
e2	z_j	$\Rightarrow$	w_i	.02	(.01)		.02	(.01)		.02	(.01)		.02	(.01)	
e3	z_j	$\Rightarrow$	w_j	.32	(.52)		.32	(.52)		.32	(.52)		.32	(.52)	
e4	z_j	$\Rightarrow$	$x_{ij}w_i$	-3.35	(4.06)		-3.35	(4.06)		-3.35	(4.06)		-3.34	(4.06)	
e5	z_j	$\Rightarrow$	$x_{ij}w_j$	-.14	(1.34)		-.14	(1.34)		-.14	(1.34)		-.14	(1.34)	
	i1		y_{ij}	50.45	(.18)	**	49.33	(.15)	**	45.94	(.11)	**	53.82	(.14)	**
	i2		m_{ij}	.01	(.15)		.01	(.15)		.01	(.15)		.01	(.15)	
	i3		x_{ij}	.01	(.16)		.01	(.16)		.01	(.16)		.01	(.16)	
	i4		w_i	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
	i5		w_j	.00	(.25)		.00	(.25)		.00	(.25)		.00	(.25)	
	i6		$x_{ij}w_i$	4.49	(1.50)	**	4.49	(1.50)	**	4.49	(1.50)	**	4.49	(1.50)	**
	i7		$x_{ij}w_j$	.33	(.51)		.33	(.51)		.33	(.51)		.33	(.51)	
	r1		y_{ij}	65.67	(1.18)	**	95.45	(2.00)	**	45.82	(1.12)	**	91.92	(1.99)	**
	r2		m_{ij}	75.61	(2.40)	**	75.61	(2.40)	**	75.60	(2.40)	**	75.61	(2.40)	**
	r3		x_{ij}	58.87	(1.60)	**	58.87	(1.60)	**	58.87	(1.60)	**	58.87	(1.60)	**
	r4		w_i	105.47	(2.65)	**	105.47	(2.65)	**	105.47	(2.65)	**	105.47	(2.65)	**
	r5		w_j	7.43	(.73)	**	7.43	(.73)	**	7.43	(.73)	**	7.43	(.73)	**
	r6		$x_{ij}w_i$	8302.65	(504.60)	**	8302.62	(504.71)	**	8302.52	(504.72)	**	8302.64	(504.71)	**
	r7		$x_{ij}w_j$	483.79	(51.38)	**	483.79	(51.38)	**	483.79	(51.38)	**	483.79	(51.38)	**
	cov2	$(x_{ij}, x_{ij}w_i)$		3.66	(1.43)	*	3.66	(1.43)	*	3.66	(1.43)	*	3.65	(1.43)	*
	cov3	$(x_{ij}, x_{ij}w_j)$		.11	(.47)		.11	(.47)		.11	(.47)		.11	(.47)	

Note * p<.05; ** p<.01; *** p<.001, $\Rightarrow$ express regression coefficient, unstandardized estimates.

Table 4A.5 Path Analysis Estimates, Thailand

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.34	(.02)	**	.30	(.02)	**	-.30	(.02)	**	.35	(.03)	**
c2	w_i	⇒	y_{ij}	.06	(.01)	**	.17	(.02)	**	.00	(.01)		.15	(.02)	**
c3	w_j	⇒	y_{ij}	.12	(.06)	*	.24	(.07)	**	-.03	(.05)		.17	(.07)	*
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	⇒	y_{ij}	.02	(.01)	*	.01	(.01)		-.01	(.01)	*	.01	(.01)	**
b1	m_{ij}	⇒	y_{ij}	-.10	(.02)	**	.15	(.02)	**	.42	(.02)	**	.14	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.46	(.02)	**	-.46	(.02)	**	-.46	(.02)	**	-.46	(.02)	**
a2	w_i	⇒	m_{ij}	.01	(.01)		.01	(.01)		.01	(.01)		.01	(.01)	
a3	w_j	⇒	m_{ij}	-.02	(.06)		-.02	(.06)		-.02	(.06)		-.02	(.06)	
a4	$x_{ij}w_i$	⇒	m_{ij}	-.01	(.00)	**	-.01	(.00)	**	-.01	(.00)	**	-.01	(.00)	**
a5	$x_{ij}w_j$	⇒	m_{ij}	-.03	(.01)	**	-.03	(.01)	**	-.03	(.01)	**	-.03	(.01)	**
d1	z_i	⇒	x_{ij}	.17	(.10)		.17	(.10)		.17	(.10)		.17	(.10)	
d2	z_i	⇒	w_i	.19	(.18)		.19	(.18)		.19	(.18)		.19	(.18)	
d3	z_i	⇒	w_j	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
d4	z_i	⇒	$x_{ij}w_i$	-.95	(1.33)		-.95	(1.33)		-.95	(1.33)		-.95	(1.33)	
d5	z_i	⇒	$x_{ij}w_j$	.17	(.39)		.17	(.39)		.17	(.39)		.17	(.39)	
e1	z_j	⇒	x_{ij}	5.11	(.52)	**	5.11	(.52)	**	5.11	(.52)	**	5.11	(.52)	**
e2	z_j	⇒	w_i	-.02	(.06)		-.02	(.06)		-.02	(.06)		-.02	(.06)	
e3	z_j	⇒	w_j	1.34	(.35)	**	1.34	(.35)	**	1.34	(.35)	**	1.34	(.35)	**
e4	z_j	⇒	$x_{ij}w_i$	-1.82	(1.14)		-1.82	(1.14)		-1.82	(1.14)		-1.82	(1.14)	
e5	z_j	⇒	$x_{ij}w_j$	8.22	(3.47)	*	8.22	(3.47)	*	8.22	(3.47)	*	8.22	(3.47)	*
	i1		y_{ij}	43.32	(.15)	**	49.21	(.16)	**	56.99	(.17)	**	45.68	(.19)	**
	i2		m_{ij}	.39	(.20)	*	.39	(.20)	*	.39	(.20)	*	.39	(.20)	*
	i3		x_{ij}	.01	(.33)		.01	(.33)		.01	(.33)		.01	(.33)	
	i4		w_i	.02	(.03)		.02	(.03)		.02	(.03)		.02	(.03)	
	i5		w_j	-.01	(.22)		-.01	(.22)		-.01	(.22)		-.01	(.22)	
	i6		$x_{ij}w_i$	10.66	(.87)	**	10.66	(.87)	**	10.66	(.87)	**	10.66	(.87)	**
	i7		$x_{ij}w_j$	10.86	(1.56)	**	10.86	(1.56)	**	10.86	(1.56)	**	10.86	(1.56)	**
	r1		y_{ij}	31.25	(1.32)	**	63.43	(1.25)	**	41.65	(1.49)	**	69.65	(1.84)	**
	r2		m_{ij}	44.94	(1.96)	**	44.94	(1.96)	**	44.94	(1.96)	**	44.94	(1.96)	**
	r3		x_{ij}	48.92	(2.77)	**	48.92	(2.77)	**	48.92	(2.77)	**	48.92	(2.77)	**
	r4		w_i	57.18	(1.98)	**	57.19	(1.99)	**	57.19	(1.98)	**	57.18	(1.98)	**
	r5		w_j	7.87	(1.07)	**	7.87	(1.07)	**	7.87	(1.07)	**	7.87	(1.07)	**
	r6		$x_{ij}w_i$	3647.29	(274.10)	**	3647.26	(274.10)	**	3647.30	(274.10)	**	3647.27	(274.10)	**
	r7		$x_{ij}w_j$	594.94	(96.29)	**	594.94	(96.29)	**	594.94	(96.29)	**	594.94	(96.29)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	10.72	(.85)	**	10.73	(.85)	**	10.72	(.85)	**	10.72	(.85)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	8.08	(1.52)	**	8.08	(1.52)	**	8.08	(1.53)	**	8.08	(1.53)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.6 Path Analysis Estimates, Taiwan

Label	Path			Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
				E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.34	(.02)	**	.22	(.02)	**	-.26	(.01)	**	.32	(.02)	**
c2	w_i	⇒	y_{ij}	.11	(.01)	**	.10	(.02)	**	-.05	(.01)	**	.02	(.02)	
c3	w_j	⇒	y_{ij}	.05	(.04)		.05	(.04)		-.10	(.04)	**	.01	(.06)	
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		-.01	(.00)	*	.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	⇒	y_{ij}	-.01	(.01)		-.01	(.01)	*	.00	(.00)		.00	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.22	(.01)	**	-.09	(.02)	**	.40	(.02)	**	-.09	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.43	(.02)	**	-.44	(.02)	**	-.44	(.02)	**	-.43	(.02)	**
a2	w_i	⇒	m_{ij}	-.02	(.02)		-.02	(.02)		-.02	(.02)		-.02	(.02)	
a3	w_j	⇒	m_{ij}	-.05	(.05)		-.05	(.05)		-.05	(.05)		-.05	(.05)	
a4	$x_{ij}w_i$	⇒	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
a5	$x_{ij}w_j$	⇒	m_{ij}	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
d1	z_i	⇒	x_{ij}	2.62	(.14)	**	2.62	(.14)	**	2.62	(.14)	**	2.62	(.14)	**
d2	z_i	⇒	w_i	.82	(.16)	**	.82	(.16)	**	.82	(.16)	**	.82	(.16)	**
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	-3.85	(1.70)	*	-3.85	(1.70)	*	-3.85	(1.70)	*	-3.85	(1.70)	*
d5	z_i	⇒	$x_{ij}w_j$	.44	(.84)		.44	(.84)		.44	(.84)		.44	(.84)	
e1	z_j	⇒	x_{ij}	6.05	(.41)	**	6.05	(.41)	**	6.05	(.41)	**	6.05	(.41)	**
e2	z_j	⇒	w_i	.00	(.03)		.00	(.03)		.00	(.03)		.00	(.03)	
e3	z_j	⇒	w_j	1.22	(.53)	*	1.22	(.53)	*	1.22	(.53)	*	1.22	(.53)	*
e4	z_j	⇒	$x_{ij}w_i$	-6.02	(2.54)	*	-6.02	(2.54)	*	-6.02	(2.54)	*	-6.02	(2.54)	*
e5	z_j	⇒	$x_{ij}w_j$	.62	(3.05)		.62	(3.05)		.62	(3.05)		.62	(3.05)	
	i1		y_{ij}	55.51	(.14)	**	56.93	(.15)	**	44.00	(.13)	**	50.22	(.20)	**
	i2		m_{ij}	.02	(.14)		.02	(.14)		.02	(.14)		.02	(.14)	
	i3		x_{ij}	.03	(.20)		.03	(.20)		.03	(.20)		.03	(.20)	
	i4		w_i	.01	(.02)		.01	(.02)		.01	(.02)		.01	(.02)	
	i5		w_j	.00	(.28)		.00	(.28)		.00	(.28)		.00	(.28)	
	i6		$x_{ij}w_i$	15.86	(1.25)	**	15.86	(1.25)	**	15.86	(1.25)	**	15.86	(1.25)	**
	i7		$x_{ij}w_j$	2.90	(1.02)	**	2.90	(1.02)	**	2.90	(1.02)	**	2.90	(1.02)	**
	r1		y_{ij}	67.40	(1.45)	**	76.28	(1.60)	**	62.99	(1.58)	**	87.44	(1.82)	**
	r2		m_{ij}	83.74	(2.35)	**	83.74	(2.35)	**	83.75	(2.35)	**	83.74	(2.35)	**
	r3		x_{ij}	73.74	(1.79)	**	73.74	(1.79)	**	73.74	(1.79)	**	73.74	(1.79)	**
	r4		w_i	90.65	(2.97)	**	90.65	(2.97)	**	90.65	(2.97)	**	90.65	(2.97)	**
	r5		w_j	10.01	(1.18)	**	10.01	(1.18)	**	10.01	(1.18)	**	10.01	(1.18)	**
	r6		$x_{ij}w_i$	9951.32	(688.61)	**	9951.35	(688.60)	**	9951.32	(688.59)	**	9951.35	(688.60)	**
	r7		$x_{ij}w_j$	880.51	(90.87)	**	880.51	(90.87)	**	880.51	(90.87)	**	880.51	(90.87)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	14.33	(1.17)	**	14.32	(1.17)	**	14.32	(1.17)	**	14.32	(1.17)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	1.09	(.76)		1.09	(.76)		1.09	(.76)		1.09	(.76)	

Note * $p<.05$; ** $p<.01$; *** $p<.001$, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.7 Path Analysis Estimates, Chile

Label	Path			Gender Equality E	(SE)		Ethnic Rights E	(SE)		Permissibility of Corruption E	(SE)		Freedom of Speech E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.34	(.03)	**	.23	(.03)	**	-.24	(.02)	**	.16	(.03)	**
c2	w_i	⇒	y_{ij}	.13	(.01)	**	.15	(.01)	**	-.01	(.02)		.10	(.02)	**
c3	w_j	⇒	y_{ij}	.23	(.05)	**	.23	(.05)	**	-.05	(.03)		.10	(.05)	
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		-.01	(.00)	**
c5	$x_{ij}w_j$	⇒	y_{ij}	.00	(.00)		-.01	(.01)		.00	(.00)		-.02	(.01)	**
b1	m_{ij}	⇒	y_{ij}	-.22	(.01)	**	-.09	(.02)	**	.52	(.02)	**	-.10	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.68	(.02)	**	-.68	(.02)	**	-.68	(.02)	**	-.68	(.02)	**
a2	w_i	⇒	m_{ij}	-.01	(.01)		-.01	(.01)		-.01	(.01)		-.01	(.01)	
a3	w_j	⇒	m_{ij}	-.10	(.04)	*	-.10	(.04)	*	-.10	(.04)	*	-.10	(.04)	*
a4	$x_{ij}w_i$	⇒	m_{ij}	-.01	(.00)	*	-.01	(.00)	*	-.01	(.00)	*	-.01	(.00)	*
a5	$x_{ij}w_j$	⇒	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d1	z_i	⇒	x_{ij}	1.32	(.19)	**	1.32	(.19)	**	1.32	(.19)	**	1.32	(.19)	**
d2	z_i	⇒	w_i	.04	(.21)		.04	(.21)		.04	(.21)		.04	(.21)	
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	.63	(2.03)		.62	(2.03)		.62	(2.04)		.60	(2.04)	
d5	z_i	⇒	$x_{ij}w_j$	.18	(.76)		.18	(.76)		.18	(.76)		.18	(.76)	
e1	z_j	⇒	x_{ij}	6.02	(.38)	**	6.02	(.38)	**	6.02	(.38)	**	6.02	(.38)	**
e2	z_j	⇒	w_i	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
e3	z_j	⇒	w_j	1.43	(.34)	**	1.43	(.34)	**	1.43	(.34)	**	1.43	(.34)	**
e4	z_j	⇒	$x_{ij}w_i$	-3.20	(1.51)	*	-3.19	(1.51)	*	-3.20	(1.51)	*	-3.18	(1.51)	*
e5	z_j	⇒	$x_{ij}w_j$	.86	(2.84)		.86	(2.84)		.86	(2.84)		.86	(2.84)	
	i1		y_{ij}	51.05	(.19)	**	54.52	(.22)	**	48.67	(.15)	**	54.27	(.18)	**
	i2		m_{ij}	.06	(.21)		.06	(.21)		.06	(.21)		.06	(.21)	
	i3		x_{ij}	.02	(.26)		.02	(.26)		.02	(.26)		.02	(.26)	
	i4		w_i	-.01	(.01)		-.01	(.01)		-.01	(.01)		-.01	(.01)	
	i5		w_j	-.01	(.30)		-.01	(.30)		-.01	(.30)		-.01	(.30)	
	i6		$x_{ij}w_i$	10.36	(1.37)	**	10.36	(1.37)	**	10.37	(1.37)	**	10.37	(1.37)	**
	i7		$x_{ij}w_j$	9.50	(1.51)	**	9.50	(1.51)	**	9.50	(1.51)	**	9.50	(1.51)	**
	r1		y_{ij}	63.67	(1.52)	**	79.11	(2.23)	**	59.66	(2.47)	**	85.50	(1.54)	**
	r2		m_{ij}	73.50	(2.83)	**	73.49	(2.83)	**	73.50	(2.83)	**	73.49	(2.83)	**
	r3		x_{ij}	58.55	(2.03)	**	58.55	(2.03)	**	58.55	(2.03)	**	58.55	(2.03)	**
	r4		w_i	86.95	(2.69)	**	86.96	(2.69)	**	86.95	(2.69)	**	86.95	(2.69)	**
	r5		w_j	14.02	(1.68)	**	14.02	(1.68)	**	14.02	(1.68)	**	14.02	(1.68)	**
	r6		$x_{ij}w_i$	6954.73	(649.39)	**	6954.77	(649.39)	**	6954.77	(649.39)	**	6954.83	(649.42)	**
	r7		$x_{ij}w_j$	1164.50	(154.67)	**	1164.50	(154.67)	**	1164.50	(154.67)	**	1164.50	(154.67)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	10.38	(1.35)	**	10.38	(1.35)	**	10.37	(1.35)	**	10.38	(1.35)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	5.44	(1.08)	**	5.44	(1.08)	**	5.43	(1.08)	**	5.44	(1.08)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.8 Path Analysis Estimates, Colombia

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.32	(.02)	**	.22	(.02)	**	-.22	(.02)	**	.20	(.04)	**
c2	w_i	⇒	y_{ij}	.07	(.01)	**	.18	(.02)	**	-.01	(.01)		.12	(.02)	**
c3	w_j	⇒	y_{ij}	.20	(.05)	**	.27	(.05)	**	-.05	(.05)		.17	(.07)	*
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		-.01	(.00)	*
c5	$x_{ij}w_j$	⇒	y_{ij}	.01	(.00)	**	.01	(.00)		.00	(.01)		-.01	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.32	(.01)	**	-.11	(.02)	**	.52	(.02)	**	-.14	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.59	(.02)	**	-.59	(.02)	**	-.59	(.02)	**	-.59	(.02)	**
a2	w_i	⇒	m_{ij}	-.03	(.02)		-.03	(.02)		-.03	(.02)		-.03	(.02)	
a3	w_j	⇒	m_{ij}	-.15	(.05)	**	-.15	(.05)	**	-.15	(.05)	**	-.15	(.05)	**
a4	$x_{ij}w_i$	⇒	m_{ij}	-.01	(.00)	**	-.01	(.00)	**	-.01	(.00)	**	-.01	(.00)	**
a5	$x_{ij}w_j$	⇒	m_{ij}	-.01	(.01)	**	-.01	(.01)	**	-.01	(.01)	**	-.01	(.01)	**
d1	z_i	⇒	x_{ij}	.95	(.16)	**	.95	(.16)	**	.95	(.16)	**	.95	(.16)	**
d2	z_i	⇒	w_i	.52	(.15)	**	.52	(.15)	**	.52	(.15)	**	.52	(.15)	**
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	.34	(1.26)		.34	(1.26)		.33	(1.26)		.34	(1.26)	
d5	z_i	⇒	$x_{ij}w_j$	.82	(.45)		.82	(.45)		.82	(.45)		.82	(.45)	
e1	z_j	⇒	x_{ij}	4.96	(.38)	**	4.96	(.38)	**	4.96	(.38)	**	4.96	(.38)	**
e2	z_j	⇒	w_i	.01	(.04)		.01	(.04)		.01	(.04)		.02	(.04)	
e3	z_j	⇒	w_j	2.44	(.41)	**	2.44	(.41)	**	2.44	(.41)	**	2.44	(.41)	**
e4	z_j	⇒	$x_{ij}w_i$	-1.05	(1.38)		-1.04	(1.38)		-1.04	(1.38)		-1.06	(1.38)	
e5	z_j	⇒	$x_{ij}w_j$	4.54	(2.97)		4.54	(2.97)		4.54	(2.97)		4.54	(2.97)	
	i1		y_{ij}	49.03	(.14)	**	52.82	(.16)	**	48.26	(.13)	**	51.28	(.20)	**
	i2		m_{ij}	.26	(.19)		.26	(.19)		.26	(.19)		.26	(.19)	
	i3		x_{ij}	.00	(.26)		.00	(.26)		.00	(.26)		.00	(.26)	
	i4		w_i	-.05	(.02)	**	-.05	(.02)	**	-.05	(.02)	**	-.05	(.02)	**
	i5		w_j	-.04	(.24)		-.04	(.24)		-.04	(.24)		-.04	(.24)	
	i6		$x_{ij}w_i$	13.55	(.89)	**	13.55	(.89)	**	13.55	(.89)	**	13.55	(.89)	**
	i7		$x_{ij}w_j$	7.63	(1.08)	**	7.63	(1.08)	**	7.63	(1.08)	**	7.63	(1.08)	**
	r1		y_{ij}	50.73	(1.31)	**	68.57	(1.33)	**	49.26	(1.50)	**	84.09	(1.48)	**
	r2		m_{ij}	58.51	(1.98)	**	58.51	(1.98)	**	58.50	(1.98)	**	58.51	(1.98)	**
	r3		x_{ij}	56.38	(1.85)	**	56.38	(1.85)	**	56.38	(1.85)	**	56.38	(1.85)	**
	r4		w_i	66.41	(2.15)	**	66.42	(2.15)	**	66.41	(2.15)	**	66.40	(2.15)	**
	r5		w_j	9.44	(1.39)	**	9.44	(1.39)	**	9.44	(1.39)	**	9.44	(1.39)	**
	r6		$x_{ij}w_i$	4266.85	(299.57)	**	4266.86	(299.56)	**	4266.86	(299.57)	**	4266.87	(299.56)	**
	r7		$x_{ij}w_j$	688.05	(80.40)	**	688.05	(80.40)	**	688.05	(80.40)	**	688.05	(80.40)	**
	cov2		(x_{ij}, $x_{ij}w_i$)	13.48	(.90)	**	13.49	(.90)	**	13.48	(.90)	**	13.46	(.90)	**
	cov3		(x_{ij}, $x_{ij}w_j$)	3.50	(.72)	**	3.50	(.72)	**	3.50	(.72)	**	3.50	(.72)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.9 Path Analysis Estimates, Dominican Republic

Label	Path			Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
				E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.35	(.03)	**	.20	(.04)	**	-.19	(.03)	**	.38	(.05)	**
c2	w_i	⇒	y_{ij}	.06	(.01)	**	.09	(.02)	**	-.01	(.01)		.09	(.02)	**
c3	w_j	⇒	y_{ij}	.08	(.05)		.06	(.08)		.00	(.06)		.12	(.08)	
c4	$x_{ij}w_i$	⇒	y_{ij}	.01	(.00)	**	.01	(.00)		-.01	(.00)	*	.01	(.00)	*
c5	$x_{ij}w_j$	⇒	y_{ij}	.02	(.01)	**	.00	(.01)		-.01	(.01)		.01	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.18	(.02)	**	.08	(.02)	**	.57	(.02)	**	.06	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.63	(.03)	**	-.63	(.03)	**	-.63	(.03)	**	-.63	(.03)	**
a2	w_i	⇒	m_{ij}	-.06	(.02)	**	-.06	(.02)	**	-.06	(.02)	**	-.06	(.02)	**
a3	w_j	⇒	m_{ij}	-.21	(.10)	*	-.21	(.10)	*	-.21	(.10)	*	-.21	(.10)	*
a4	$x_{ij}w_i$	⇒	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
a5	$x_{ij}w_j$	⇒	m_{ij}	-.03	(.01)	**	-.04	(.01)	**	-.03	(.01)	**	-.03	(.01)	**
d1	z_i	⇒	x_{ij}	.33	(.18)		.33	(.18)		.33	(.18)		.33	(.18)	
d2	z_i	⇒	w_i	.09	(.25)		.08	(.25)		.08	(.25)		.09	(.24)	
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	2.29	(1.62)		2.26	(1.62)		2.27	(1.62)		2.30	(1.62)	
d5	z_i	⇒	$x_{ij}w_j$	-.19	(.39)		-.19	(.39)		-.19	(.39)		-.19	(.39)	
e1	z_j	⇒	x_{ij}	3.19	(.52)	**	3.19	(.52)	**	3.19	(.52)	**	3.19	(.52)	**
e2	z_j	⇒	w_i	.06	(.05)		.05	(.05)		.05	(.05)		.05	(.04)	
e3	z_j	⇒	w_j	2.14	(.45)	**	2.14	(.45)	**	2.14	(.45)	**	2.14	(.45)	**
e4	z_j	⇒	$x_{ij}w_i$	4.01	(2.70)		3.96	(2.71)		3.96	(2.70)		3.96	(2.71)	
e5	z_j	⇒	$x_{ij}w_j$	3.72	(2.25)		3.72	(2.25)		3.72	(2.25)		3.72	(2.25)	
	i1		y_{ij}	42.97	(.19)	**	50.69	(.26)	**	55.19	(.17)	**	48.78	(.29)	**
	i2		m_{ij}	.30	(.31)		.29	(.31)		.29	(.31)		.30	(.31)	
	i3		x_{ij}	.05	(.24)		.05	(.24)		.05	(.24)		.05	(.24)	
	i4		w_i	-.09	(.05)	*	-.09	(.05)	*	-.09	(.05)		-.08	(.05)	
	i5		w_j	-.05	(.25)		-.05	(.25)		-.05	(.25)		-.05	(.25)	
	i6		$x_{ij}w_i$	13.85	(1.38)	**	13.86	(1.37)	**	13.88	(1.38)	**	13.91	(1.38)	**
	i7		$x_{ij}w_j$	5.30	(1.01)	**	5.30	(1.01)	**	5.30	(1.01)	**	5.30	(1.01)	**
	r1		y_{ij}	39.72	(2.14)	**	75.80	(1.59)	**	41.81	(2.18)	**	91.12	(2.29)	**
	r2		m_{ij}	79.75	(3.66)	**	79.75	(3.66)	**	79.68	(3.65)	**	79.76	(3.66)	**
	r3		x_{ij}	41.42	(1.35)	**	41.42	(1.35)	**	41.42	(1.35)	**	41.42	(1.35)	**
	r4		w_i	89.49	(2.94)	**	89.51	(2.95)	**	89.50	(2.95)	**	89.45	(2.94)	**
	r5		w_j	8.26	(1.03)	**	8.26	(1.03)	**	8.26	(1.03)	**	8.26	(1.03)	**
	r6		$x_{ij}w_i$	4125.33	(383.56)	**	4125.24	(383.59)	**	4125.19	(383.57)	**	4125.13	(383.59)	**
	r7		$x_{ij}w_j$	425.20	(58.35)	**	425.20	(58.35)	**	425.20	(58.35)	**	425.20	(58.35)	**
	cov2	$(x_{ij}, x_{ij}w_i)$		13.63	(1.36)	**	13.67	(1.36)	**	13.65	(1.36)	**	13.60	(1.37)	**
	cov3	$(x_{ij}, x_{ij}w_j)$		3.60	(.87)	**	3.60	(.87)	**	3.60	(.87)	**	3.59	(.87)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.10 Path Analysis Estimates, Guatemala

Label	Path			Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
				E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.37	(.03)	**	.26	(.04)	**	-.37	(.03)	**	.24	(.03)	**
c2	w_i	⇒	y_{ij}	.08	(.01)	**	.11	(.02)	**	-.07	(.01)	**	.12	(.02)	**
c3	w_j	⇒	y_{ij}	.16	(.06)	*	.21	(.07)	**	.00	(.06)		.17	(.05)	**
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)	*	-.01	(.00)	**
c5	$x_{ij}w_j$	⇒	y_{ij}	.02	(.01)	**	.01	(.01)		-.01	(.01)		-.01	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.29	(.02)	**	-.09	(.02)	**	.51	(.02)	**	-.06	(.02)	*
a1	x_{ij}	⇒	m_{ij}	-.62	(.03)	**	-.62	(.03)	**	-.62	(.03)	**	-.62	(.03)	**
a2	w_i	⇒	m_{ij}	-.06	(.02)	**	-.06	(.02)	**	-.06	(.02)	**	-.06	(.02)	**
a3	w_j	⇒	m_{ij}	-.09	(.06)		-.09	(.06)		-.09	(.06)		-.09	(.06)	
a4	$x_{ij}w_i$	⇒	m_{ij}	-.01	(.00)	*	-.01	(.00)	*	-.01	(.00)	*	-.01	(.00)	*
a5	$x_{ij}w_j$	⇒	m_{ij}	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**
d1	z_i	⇒	x_{ij}	.81	(.14)	**	.81	(.14)	**	.81	(.14)	**	.81	(.14)	**
d2	z_i	⇒	w_i	.06	(.24)		.07	(.24)		.07	(.24)		.07	(.24)	
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	-.83	(1.76)		-.83	(1.76)		-.82	(1.76)		-.81	(1.76)	
d5	z_i	⇒	$x_{ij}w_j$	.19	(.55)		.19	(.55)		.19	(.55)		.19	(.55)	
e1	z_j	⇒	x_{ij}	6.32	(.42)	**	6.32	(.42)	**	6.32	(.42)	**	6.32	(.42)	**
e2	z_j	⇒	w_i	.04	(.04)		.05	(.05)		.04	(.04)		.04	(.05)	
e3	z_j	⇒	w_j	2.12	(.48)	**	2.12	(.48)	**	2.12	(.48)	**	2.12	(.48)	**
e4	z_j	⇒	$x_{ij}w_i$	-.90	(1.79)		-.89	(1.79)		-.90	(1.78)		-.90	(1.78)	
e5	z_j	⇒	$x_{ij}w_j$	.36	(4.33)		.36	(4.33)		.36	(4.33)		.36	(4.33)	
	i1		y_{ij}	48.69	(.18)	**	54.93	(.21)	**	50.10	(.19)	**	50.18	(.17)	**
	i2		m_{ij}	.22	(.22)		.22	(.22)		.22	(.22)		.22	(.22)	
	i3		x_{ij}	.00	(.25)		.00	(.25)		.00	(.25)		.00	(.25)	
	i4		w_i	-.02	(.02)		-.02	(.02)		-.02	(.02)		-.02	(.02)	
	i5		w_j	-.02	(.24)		-.02	(.24)		-.02	(.24)		-.02	(.24)	
	i6		$x_{ij}w_i$	10.05	(1.18)	**	10.06	(1.18)	**	10.05	(1.18)	**	10.06	(1.18)	**
	i7		$x_{ij}w_j$	7.55	(1.26)	**	7.55	(1.26)	**	7.55	(1.26)	**	7.55	(1.26)	**
	r1		y_{ij}	49.20	(1.23)	**	64.09	(1.53)	**	46.57	(2.24)	**	79.19	(2.01)	**
	r2		m_{ij}	54.51	(2.51)	**	54.50	(2.51)	**	54.50	(2.51)	**	54.50	(2.51)	**
	r3		x_{ij}	42.11	(1.46)	**	42.11	(1.46)	**	42.11	(1.46)	**	42.11	(1.46)	**
	r4		w_i	81.86	(2.36)	**	81.85	(2.36)	**	81.86	(2.36)	**	81.88	(2.37)	**
	r5		w_j	8.23	(.98)	**	8.23	(.98)	**	8.23	(.98)	**	8.23	(.98)	**
	r6		$x_{ij}w_i$	4659.09	(403.12)	**	4659.12	(403.13)	**	4659.10	(403.12)	**	4659.03	(403.10)	**
	r7		$x_{ij}w_j$	621.65	(106.24)	**	621.65	(106.24)	**	621.65	(106.24)	**	621.65	(106.24)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	10.08	(1.20)	**	10.06	(1.20)	**	10.09	(1.19)	**	10.10	(1.19)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	2.62	(.78)	**	2.62	(.78)	**	2.62	(.78)	**	2.62	(.78)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.11 Path Analysis Estimates, Mexico

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.26	(.01)	**	.29	(.02)	**	-.22	(.02)	**	.29	(.02)	**
c2	w_i	⇒	y_{ij}	.05	(.01)	**	.17	(.01)	**	-.02	(.01)		.11	(.02)	**
c3	w_j	⇒	y_{ij}	.06	(.03)	*	.31	(.05)	**	-.01	(.04)		.25	(.05)	**
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		-.01	(.00)	**	.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	⇒	y_{ij}	.00	(.00)		-.01	(.01)		.01	(.00)		-.01	(.01)	*
b1	m_{ij}	⇒	y_{ij}	-.16	(.01)	**	-.08	(.02)	**	.57	(.02)	**	-.10	(.02)	**
a1	x_{ij}	⇒	m_{ij}	-.75	(.02)	**	-.75	(.02)	**	-.75	(.02)	**	-.75	(.02)	**
a2	w_i	⇒	m_{ij}	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
a3	w_j	⇒	m_{ij}	-.05	(.08)		-.05	(.08)		-.05	(.08)		-.05	(.08)	
a4	$x_{ij}w_i$	⇒	m_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
a5	$x_{ij}w_j$	⇒	m_{ij}	-.02	(.01)	*	-.02	(.01)	*	-.02	(.01)	*	-.02	(.01)	*
d1	z_i	⇒	x_{ij}	.84	(.14)	**	.84	(.14)	**	.84	(.14)	**	.84	(.14)	**
d2	z_i	⇒	w_i	.19	(.12)		.19	(.12)		.19	(.12)		.19	(.12)	
d3	z_i	⇒	w_j	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
d4	z_i	⇒	$x_{ij}w_i$	.50	(1.39)		.50	(1.39)		.50	(1.39)		.50	(1.39)	
d5	z_i	⇒	$x_{ij}w_j$	.68	(.57)		.68	(.57)		.68	(.57)		.68	(.57)	
e1	z_j	⇒	x_{ij}	5.13	(.33)	**	5.13	(.33)	**	5.13	(.33)	**	5.13	(.33)	**
e2	z_j	⇒	w_i	.01	(.02)		.02	(.02)		.01	(.02)		.01	(.02)	
e3	z_j	⇒	w_j	1.30	(.32)	**	1.30	(.32)	**	1.30	(.32)	**	1.30	(.32)	**
e4	z_j	⇒	$x_{ij}w_i$	.31	(2.09)		.30	(2.09)		.30	(2.09)		.30	(2.09)	
e5	z_j	⇒	$x_{ij}w_j$	2.77	(2.75)		2.77	(2.75)		2.77	(2.75)		2.77	(2.75)	
	i1		y_{ij}	45.36	(.08)	**	52.41	(.15)	**	49.22	(.13)	**	50.17	(.19)	**
	i2		m_{ij}	.19	(.28)		.19	(.28)		.19	(.28)		.19	(.28)	
	i3		x_{ij}	.01	(.21)		.01	(.21)		.01	(.21)		.01	(.21)	
	i4		w_i	-.02	(.01)		-.02	(.01)	*	-.02	(.01)		-.02	(.01)	*
	i5		w_j	-.02	(.21)		-.02	(.21)		-.02	(.21)		-.02	(.21)	
	i6		$x_{ij}w_i$	10.19	(.95)	**	10.20	(.95)	**	10.19	(.95)	**	10.19	(.95)	**
	i7		$x_{ij}w_j$	5.92	(1.22)	**	5.92	(1.22)	**	5.92	(1.22)	**	5.92	(1.22)	**
	r1		y_{ij}	29.76	(.83)	**	74.50	(1.49)	**	56.26	(2.01)	**	91.26	(1.81)	**
	r2		m_{ij}	82.02	(2.69)	**	82.02	(2.69)	**	82.01	(2.69)	**	82.02	(2.69)	**
	r3		x_{ij}	57.81	(1.84)	**	57.81	(1.84)	**	57.81	(1.84)	**	57.81	(1.84)	**
	r4		w_i	77.82	(2.10)	**	77.82	(2.10)	**	77.82	(2.10)	**	77.82	(2.10)	**
	r5		w_j	9.00	(.86)	**	9.00	(.86)	**	9.00	(.86)	**	9.00	(.86)	**
	r6		$x_{ij}w_i$	4957.94	(278.78)	**	4957.96	(278.78)	**	4957.94	(278.78)	**	4957.94	(278.77)	**
	r7		$x_{ij}w_j$	702.12	(74.99)	**	702.12	(74.99)	**	702.12	(74.99)	**	702.12	(74.99)	**
	cov2		$(x_{ij}, x_{ij}w_i)$	10.14	(.93)	**	10.15	(.93)	**	10.15	(.93)	**	10.14	(.93)	**
	cov3		$(x_{ij}, x_{ij}w_j)$	3.36	(.75)	**	3.36	(.75)	**	3.36	(.75)	**	3.36	(.75)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

Table 4A.12 Path Analysis Estimates, Paraguay

				Gender Equality			Ethnic Rights			Permissibility of Corruption			Freedom of Speech		
Label	Path			E	(SE)		E	(SE)		E	(SE)		E	(SE)	
c1	x_{ij}	⇒	y_{ij}	.33	(.03)	**	.14	(.03)	**	-.18	(.03)	**	.23	(.03)	**
c2	w_i	⇒	y_{ij}	.10	(.02)	**	.14	(.02)	**	-.03	(.02)		.11	(.03)	**
c3	w_j	⇒	y_{ij}	.19	(.05)	**	.12	(.07)		-.12	(.07)		.10	(.10)	
c4	$x_{ij}w_i$	⇒	y_{ij}	.00	(.00)		.00	(.00)		.00	(.00)		.00	(.00)	
c5	$x_{ij}w_j$	⇒	y_{ij}	.01	(.01)		.01	(.01)		.01	(.01)		.00	(.01)	
b1	m_{ij}	⇒	y_{ij}	-.27	(.02)	**	-.10	(.02)	**	.63	(.02)	**	-.10	(.03)	**
a1	x_{ij}	⇒	m_{ij}	-.53	(.02)	**	-.53	(.02)	**	-.53	(.02)	**	-.53	(.02)	**
a2	w_i	⇒	m_{ij}	-.05	(.02)	**	-.05	(.02)	**	-.05	(.02)	**	-.05	(.02)	**
a3	w_j	⇒	m_{ij}	-.03	(.06)		-.03	(.06)		-.03	(.06)		-.03	(.06)	
a4	$x_{ij}w_i$	⇒	m_{ij}	.00	(.00)	*	.00	(.00)	*	.00	(.00)	*	.00	(.00)	*
a5	$x_{ij}w_j$	⇒	m_{ij}	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**	-.02	(.01)	**
d1	z_i	⇒	x_{ij}	1.12	(.21)	**	1.12	(.21)	**	1.12	(.21)	**	1.12	(.21)	**
d2	z_i	⇒	w_i	.27	(.21)		.26	(.21)		.27	(.21)		.27	(.21)	
d3	z_i	⇒	w_j	.00	(.01)		.00	(.01)		.00	(.01)		.00	(.01)	
d4	z_i	⇒	$x_{ij}w_i$	1.78	(2.29)		1.78	(2.29)		1.78	(2.29)		1.78	(2.29)	
d5	z_i	⇒	$x_{ij}w_j$	-.07	(.73)		-.07	(.73)		-.07	(.73)		-.07	(.73)	
e1	z_j	⇒	x_{ij}	7.06	(.43)	**	7.06	(.43)	**	7.06	(.43)	**	7.06	(.43)	**
e2	z_j	⇒	w_i	-.01	(.09)		.00	(.09)		-.01	(.09)		-.01	(.09)	
e3	z_j	⇒	w_j	3.09	(.33)	**	3.09	(.33)	**	3.09	(.33)	**	3.09	(.33)	**
e4	z_j	⇒	$x_{ij}w_i$	-2.03	(2.06)		-2.04	(2.07)		-2.05	(2.07)		-2.06	(2.07)	
e5	z_j	⇒	$x_{ij}w_j$	9.16	(3.40)	**	9.16	(3.40)	**	9.16	(3.40)	**	9.16	(3.40)	**
	i1		y_{ij}	48.03	(.18)	**	51.42	(.23)	**	49.56	(.22)	**	51.55	(.30)	**
	i2		m_{ij}	.35	(.20)		.34	(.20)		.34	(.20)		.34	(.20)	
	i3		x_{ij}	.06	(.28)		.06	(.28)		.06	(.28)		.06	(.28)	
	i4		w_i	-.03	(.03)		-.03	(.03)		-.03	(.03)		-.03	(.03)	
	i5		w_j	-.10	(.23)		-.10	(.23)		-.10	(.23)		-.10	(.23)	
	i6		$x_{ij}w_i$	9.63	(1.19)	**	9.63	(1.19)	**	9.64	(1.19)	**	9.64	(1.19)	**
	i7		$x_{ij}w_j$	11.32	(1.57)	**	11.32	(1.57)	**	11.32	(1.57)	**	11.32	(1.57)	**
	r1		y_{ij}	51.79	(1.76)	**	63.37	(2.39)	**	47.17	(2.05)	**	90.88	(2.20)	**
	r2		m_{ij}	56.04	(2.39)	**	56.04	(2.39)	**	56.03	(2.38)	**	56.03	(2.38)	**
	r3		x_{ij}	57.74	(2.63)	**	57.74	(2.63)	**	57.74	(2.63)	**	57.74	(2.63)	**
	r4		w_i	62.07	(2.47)	**	62.07	(2.47)	**	62.07	(2.47)	**	62.07	(2.47)	**
	r5		w_j	7.39	(1.04)	**	7.39	(1.04)	**	7.39	(1.04)	**	7.39	(1.04)	**
	r6		$x_{ij}w_i$	4705.77	(407.67)	**	4705.85	(407.69)	**	4705.84	(407.66)	**	4705.85	(407.69)	**
	r7		$x_{ij}w_j$	931.48	(157.23)	**	931.48	(157.23)	**	931.48	(157.23)	**	931.48	(157.23)	**
	cov2	(x_{ij}, $x_{ij}w_i$)		9.55	(1.16)	**	9.54	(1.16)	**	9.55	(1.16)	**	9.53	(1.16)	**
	cov3	(x_{ij}, $x_{ij}w_j$)		2.57	(.89)	**	2.57	(.89)	**	2.56	(.89)	**	2.56	(.89)	**

Note * p<.05; ** p<.01; *** p<.001, ⇒ express regression coefficient, unstandardized estimates.

F) Measures

Socio Economic Status (NISB)

The National Index of Socio Economic Background (NISB) is a standardized measure at each country, with a mean of 0 and a standard deviation of 1. It averages results from three other variables present In the ICCS 2009: Books at home (HOMELIT), highest occupational status of parents (HISEI) and highest educational level of parents in approximate years of education (PAREDYRS).

Civic Knowledge (PV1CIV- PV5CIV)

Students were assessed on 80 items measuring civic and citizenship knowledge, analysis and reasoning. Items were assigned to seven booklets of 30 to 37 items (i.e. test forms, each of which contain 3 out of 7 item-clusters) following a balanced rotated design (Schulz, Ainley, & Fraillon, 2011). Students' scores were estimated via Rasch Model (Rasch, 1993), and estimates were scaled to have an international mean of 500 and a standard deviation of 100 for equally weighted countries. Plausible values were also computed for students, which only had the questionnaire data; five plausible values were estimated for analysis to overcome sampling and measurement error. In the present research, this measure was divided by 10 in order to change its scale to make the raw estimates more interpretable, therefore displaying a mean of 50 and standard deviation of 10, and afterwards centered to the grand mean.

This scale permits us to differentiate students along civic knowledge. At the lowest end, students are able to justify voluntary voting in the context of freedom of political expression, and identify that democratic leaders should be aware of the needs of the people over whom they have authority. In contrast, at the highest end, a higher civic knowledge permits students to evaluate a policy with respect to equality, to justify the separation of powers between the judiciary and parliament, and understand the reasons to regulate mass media to ensure openness for press. In short, this is a rich full measure of political sophistication.

Open Classroom Discussion (OPDISC)

Students indicated if teachers encourage and foster discussion within classrooms. It is an IRT WLE scale, with a mean of 50 and standard deviation of 10. It consists of six Likert type items with four levels for agreement. The items included in this scale are:

When discussing political and social issues during regular lessons, how often do the following things happen?

- Teachers encourage students to make up their own minds
- Teachers encourage students to express their opinions
- Students bring up current political events for discussion in class
- Students express opinions in class even when their opinions are different from most of the other students
- Teachers encourage students to discuss the issues with people having different opinions
- Teachers present several sides of the issues when explaining them in class

Authoritarianism

The measure of authoritarianism is different between the samples of Latin American countries, and those of the Asian region. In the following section, each item included as authoritarianism is described.

Latin American samples (AUTGOV): nine Likert type items were used to construct a IRT WLE scale, with four levels of agreement. Students indicated their level of agreement to these affirmations:

- It is better for government leaders to make decisions without consulting anybody.
- People in government must enforce their authority even if it means violating the rights of some citizens.
- People in government lose part of their authority when they admit their mistakes.

- People whose opinions are different than those of the government must be considered its enemies.
- The most important opinion of a country should be that of the president.
- It is fair that the government does not comply with the law when it thinks it is not necessary.
- Concentration of power in one person guarantees order.
- The government should close communication media that are critical.
- If the president does not agree with <Congress>, he/she should dissolve it

Scores were scaled with a mean of 50 and standard deviation of 10 for equally weighted countries. This measure is referred to as authoritarianism for short.

Asian samples (UNDEMGOV): Seven Likert type items were used to construct a IRT WLE scale, with four levels of agreement. Students rated their level of agreement to affirmation such as:

- As long as everyone can enjoy prosperity, it does not matter whether the government is democratic or not.
- As long as the government represents citizens' ideas it does not matter whether the government is democratic or not.
- It is acceptable for the government to act undemocratically in order to do its job more efficiently.
- The more power the government has, the more likely it is to solve its people's problems.
- It is acceptable for the government to break the law when it considers it necessary.

Scores were scaled with a mean of 50 and standard deviation of 10 for equally weighted countries. This measure is referred to as authoritarianism for short.

Gender Equality (GENEQL)

Students rated their level of agreement with six items referring to equality between men and women. Their answers were subjected to a Rasch Model, via an IRT WLE scale, with a mean of 50 and standard deviation of 10. The items included in this scale are:

- Men and women should have equal opportunities to take part in government
- Men and women should have the same rights in every way
- Women should stay out of politics
- When there are not many jobs available, men should have more right to a job than women
- Men and women should get equal pay when they are doing the same jobs
- Men are better qualified to be political leaders than women

Ethnic Rights (ETHRGT)

Students indicated their level of agreement with five items referring to equal rights to all ethnicities. Their answers were subjected to a Rasch Model, via an IRT WLE scale, with a mean of 50, and standard deviation of 10. The items included in this scale are:

- All <ethnic/racial groups> should have an equal chance to get a good education in <country of test>
- All <ethnic/racial groups> should have an equal chance to get good jobs in <country of test>
- Schools should teach students to respect members of all <ethnic/racial groups>
- <Members of all ethnic/racial groups> should be encouraged to run in elections for political office
- <Members of all ethnic/racial groups> should have the same rights

Permissibility of Corruption

The measure of Permissibility of Corruption is different between the samples of Latin American countries, and those of the Asian region. In the following section, the scales used in each region are described.

Latin American samples (ATTCORR): This was measured with five Likert type items, with a four level scale from Strongly Agree to Strongly Disagree. Students' answers were subjected to a Rasch Model with an IRT WLE scale, with a mean of 50 and standard deviation of 10. The items included in this scale are:

- It is acceptable for a civil servant to accept bribes if his salary is too low.
- It is acceptable for a civil servant to use the resources of the institution in which he/she works for personal benefit
- Good candidates grant personal benefits to voters in return for their votes.
- Paying an additional amount to a civil servant in order to obtain a personal benefit is acceptable.
- It is acceptable that a civil servant helps his/her friends by giving them employment in his/her office.
- Since public resources belong to everyone, it is acceptable that those who can keep part of them

Asian samples (CORRPUB): This was measured with three Likert type items, with a four level scale from Strongly Agree to Strongly Disagree. Student's answers were subjected to a Rasch Model with an IRT WLE scale, with a mean of 50 and standard deviation of 10. The items included in this scale are:

- It is acceptable to bribe government officials to get things done effectively.
- It doesn't matter if a public official uses resources from the institution where he/she works for his/her personal benefit.
- Preventing corruption is adults' business, it has nothing to do with me.

Freedom of Speech (DEMVAL)

Students rated their level of agreement with five items referring to freedom of speech under democracy. Their answers were subjected to a Rasch Model, via an IRT WLE scale, with a mean of 50 and standard deviation of 10. The items included in this scale are:

- Everyone should always have the right to express their opinions freely
- All people should have their social and political rights respected
- People should always be free to criticise the government publicly
- All citizens should have the right to elect their leaders freely
- People should be able to protest if they believe a law is unfair

G) Regression Estimates and Correlations between covariates and outcomes

Table 4A.13 Regression Estimates and correlations for each outcome (Asian countries)

Asian Samples		Gender Equality				Ethnic Rights				Permissibility of Corruption				Freedom of Speech			
		B		r		B		r		B		r		B		r	
Hong Kong SAR	SES	.01		.03		.03		.05		.06	**	.03		.04	*	.07	**
	Civ. Knowledge	.24	**	.31	**	.20	**	.26	**	-.24	**	-.35	**	.19	**	.25	**
	Authoritarian.	-.15	**	-.23	**	-.07	**	-.13	**	.35	**	.42	**	-.12	**	-.17	**
	Op. Class. Disc.	.11	**	.18	**	.16	**	.22	**	-.06	**	-.16	**	.08	**	.15	**
Indonesia	SES	.04		.13	**	.07	**	.12	**	-.02		-.13	**	.08	**	.14	**
	Civ. Knowledge	.31	**	.37	**	.27	**	.26	**	-.29	**	-.43	**	.29	**	.31	**
	Authoritarian.	-.07	**	-.19	**	.17	**	.05		.32	**	.43	**	.11	**	-.02	
	Op. Class. Disc.	.11	**	.20	**	.16	**	.22	**	-.09	**	-.20	**	.14	**	.21	**
Republic of Korea	SES	-.05	**	.06	**	.03		.11	**	.02		-.09	**	.02		.16	**
	Civ. Knowledge	.28	**	.33	**	.22	**	.28	**	-.19	**	-.34	**	.38	**	.45	**
	Authoritarian.	-.19	**	-.28	**	-.13	**	-.21	**	.45	**	.52	**	-.16	**	-.29	**
	Op. Class. Disc.	.06	**	.07	**	.07	**	.08	**	-.01		-.02		.04	**	.06	**
Thailand	SES	.07	**	.20	**	.09	**	.15	**	-.01		-.14	**	.06	**	.14	**
	Civ. Knowledge	.39	**	.49	**	.25	**	.28	**	-.29	**	-.47	**	.29	**	.30	**
	Authoritarian.	-.12	**	-.32	**	.13	**	-.02		.40	**	.54	**	.11	**	-.05	*
	Op. Class. Disc.	.07	**	.23	**	.17	**	.25	**	.00		-.15	**	.13	**	.23	**
Taiwan	SES	-.03	*	.14	**	.00		.10	**	.04	**	-.11	**	.04	**	.16	**
	Civ. Knowledge	.35	**	.46	**	.23	**	.29	**	-.27	**	-.44	**	.29	**	.35	**
	Authoritarian.	-.23	**	-.38	**	-.10	**	-.20	**	.41	**	.52	**	-.10	**	-.22	**
	Op. Class. Disc.	.11	**	.20	**	.11	**	.17	**	-.06	**	-.15	**	.02		.09	**

** p<.01, * p<.05.

Table 4A.14 Regression Estimates and correlations for each outcome (Latin countries)

		Gender Equality		Ethnic Rights		Permissibility of Corruption		Freedom of Speech	
Latin Samples		B	r	B	r	B	r	B	r
Chile	SES	.01	.20 **	-.02	.11 **	.02	-.19 **	-.01	.08 **
	Civ. Knowledge	.31 **	.49 **	.22 **	.31 **	-.22 **	-.52 **	.15 **	.23 **
	Authoritarian.	-.25 **	-.45 **	-.10 **	-.25 **	.53 **	.65 **	-.10 **	-.20 **
	Op. Class. Disc.	.15 **	.25 **	.18 **	.24 **	-.02	-.14 **	.12 **	.16 **
Colombia	SES	.05 **	.20 **	.03	.12 **	.03 **	-.11 **	.03	.11 **
	Civ. Knowledge	.28 **	.51 **	.19 **	.32 **	-.21 **	-.48 **	.17 **	.28 **
	Authoritarian.	-.33 **	-.52 **	-.11 **	-.26 **	.52 **	.63 **	-.12 **	-.24 **
	Op. Class. Disc.	.09 **	.25 **	.19 **	.27 **	-.02	-.18 **	.11 **	.19 **
Dominican Republic	SES	.05 **	.14 **	-.01	.00	.00	-.09 **	.03	.08 **
	Civ. Knowledge	.32 **	.46 **	.14 **	.14 **	-.15 **	-.43 **	.26 **	.26 **
	Authoritarian.	-.24 **	-.41 **	.08 **	.00	.63 **	.69 **	.04	-.07 **
	Op. Class. Disc.	.08 **	.23 **	.10 **	.13 **	-.01	-.17 **	.09 **	.16 **
Guatemala	SES	.03	.21 **	.00	.12 **	.04 *	-.18 **	.09 **	.17 **
	Civ. Knowledge	.31 **	.51 **	.24 **	.32 **	-.30 **	-.56 **	.17 **	.26 **
	Authoritarian.	-.30 **	-.49 **	-.09 **	-.25 **	.48 **	.65 **	-.05 *	-.19 **
	Op. Class. Disc.	.10 **	.23 **	.14 **	.21 **	-.06 **	-.23 **	.13 **	.19 **
Mexico	SES	-.03	.13 **	.04 **	.15 **	.00	-.16 **	.05 **	.15 **
	Civ. Knowledge	.34 **	.50 **	.25 **	.35 **	-.17 **	-.51 **	.23 **	.32 **
	Authoritarian.	-.27 **	-.47 **	-.09 **	-.26 **	.59 **	.69 **	-.10 **	-.25 **
	Op. Class. Disc.	.07 **	.17 **	.18 **	.25 **	-.02	-.13 **	.11 **	.18 **
Paraguay	SES	.05 **	.26 **	-.01	.09 **	-.03	-.23 **	.01	.12 **
	Civ. Knowledge	.31 **	.52 **	.14 **	.24 **	-.16 **	-.49 **	.20 **	.28 **
	Authoritarian.	-.28 **	-.48 **	-.12 **	-.22 **	.58 **	.68 **	-.09 **	-.22 **
	Op. Class. Disc.	.11 **	.26 **	.15 **	.21 **	-.03	-.19 **	.10 **	.18 **

** p<.01, * p<.05.